

Segmentación lingüística del color mediante un modelo Bayesiano jerárquico.

por

Teresa Efigenia Alarcón Martínez

Sometida a revisión al Departamento de Ciencias de la Computación
en el cumplimiento parcial de los requisitos para obtener el grado de

Doctor en Ciencias de la Computación

en el

CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS A.C.,

Marzo 2007

© Centro de Investigación en Matemáticas A.C., 2007

Firma del autor
Departamento de Ciencias de la Computación
Marzo de 2007

Aceptado por
Dr. Mariano Rivera Meraz
Coordinador de la Maestría en Ciencias de la Computación

Segmentación lingüística del color mediante un modelo Bayesiano jerárquico.

por

Teresa Efigenia Alarcón Martínez

Sometida a revisión al Departamento de Ciencias de la Computación
en Marzo de 2007, en el cumplimiento parcial de los
requisitos para obtener el grado de
Doctor en Ciencias de la Computación

Resumen

En este trabajo se realiza la fusión de técnicas Bayesianas de interacción espacial con un modelo de categorización del color para dar lugar a la segmentación lingüística. El modelo de categorización del color está basado en las 11 categorías universales de color propuestas por Berlin y Kay en 1969. La modelación se realiza en los espacios RGB, $L^*a^*b^*$ y $L^*u^*v^*$, usando datos obtenidos a través de experimentos de nombramiento del color con observadores humanos. La verosimilitud se modela como una combinación lineal de splines cuadráticos. Como resultado, cada voxel en una representación de color, es descrito por un vector de probabilidades, cuyas componentes expresan el grado de pertenencia a cada una de las categorías de color. Esto da lugar a un diccionario probabilístico que es usado para la segmentación a través de Campos Aleatorios de Medida Markovianos Gaussianos con entropía controlada (EC-GMMF), modelo propuesto por M. Rivera y otros en el 2005. En el trabajo de tesis se propone una generalización del modelo mencionado, que permite considerar las interacciones de tipo perceptual, que existen entre los colores básicos y que fueron establecidas por Boynton y Olso en 1987. El enfoque de segmentación que se propone simula la incertidumbre presente en zonas oscuras de color y la composición en color de una imagen, se describe en lenguaje similar al observador humano. Como producto de la segmentación se obtiene además un mapa de bordes, que se corresponde con las fronteras de los objetos más significativos, desde el punto de vista perceptual. Se propone usar el método elaborado para filtrado de imágenes en colores y en la etapa de cuantificación de algoritmos de compresión.

Abstract

In this work we combine Bayesian techniques with a color categorization model, which leads to a method for the linguistic segmentation of color images. The categorization model considers the eleven universal color categories proposed by Berlin and Kay in 1969. The modeling process is carried out in RGB, $L^*u^*v^*$ and $L^*a^*b^*$ color spaces, using data from a naming color experiment. The likelihood for each category is represented by a linear combination of quadratic splines, and as a result, each voxel in the color space is described as a vector of probabilities, whose components express the degree to which the voxel belongs to a given color category. This gives rise to a probabilistic dictionary which is used for the segmentation, in which prior spatial granularity constraints are incorporated via an entropy controlled Gauss-Markov Measure Field (EC-GMMF) model, as proposed by Rivera et. al; 2005. We give a generalization of EC-GMMF that allows one to consider the perceptual interactions between the basic colors that were experimentally established by Boynton and Olson in 1987. The new proposal considers the uncertainty in dark color regions and the color image composition is described in a similar way as human. The segmentation method allows one to obtain an edge map which may be used in a filtering process that preserves the perceptually salient borders in an image. The new proposal is also proposed for the quantization step in image color compression.

Dedicatorias

A mis padres
A mi hermana
A mi Diana.

Quiero expresar en estas letras mis agradecimientos a todos mis compañeros cubanos en el instituto de Cibernética, Matemática y Física. Muchas gracias al Dr. Roberto Rodríguez por haberme introducido el campo del procesamiento digital de imágenes.

Mis agradecimientos a todo el colectivo del CIMAT por haberme aceptado y apoyado en todo momento. Un especial agradecimiento al Dr. José Luis Marroquín, quien guió en todo momento mi trabajo, encaminó con sabiduría mi investigación y siempre me apoyó. Gracias a todos mis profesores en el programa de Doctorado: Dr. Mariano Rivera, Dr. Horebeek, a la profesora Stephanie Dunbar gracias por su apoyo y trabajo dedicado, al Dr. Arturo Hernández y al Dr. Salvador Botello. Gracias al Dr. Raúl Felipe Parada por su ayuda como maestro y como compañero. Gracias a la dirección de apoyo académico: a Lourdes, Toquina, Juanita, Lolita, Jannet, Paola, a José Castro, a Omar, al Dr. Adolfo Sánchez. Gracias a los compañeros de la biblioteca: a Marisela, a Licha, a Manuel, a Ricardo, a Liliana y al colectivo de recursos humanos y contabilidad: a Irma, Rafael, Lucha, Irma, Rafael, Adriana y Agustín.

Muchas muchas gracias a todos mis compañeros: a Oscar Dalmau, a Edgar Arce, a Omar Tzec, a María de la Paz, a Matilde y Richard, a Alonso, a Alfonso y su esposa, a Marco Bertani, a Fernando, a Joaquín, a Laura, a Lenin, a Isabel, a Patty, a Maria del Carmen, a David, a Marisa, a Angel, a Everardo (Cacho), a Juan y a Patricia Eugenia. A todos los que compartieron conmigo estos 4 años y 6 meses, gracias por todo lo recibido. Finalmente mi agradecimiento al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo dado a través del proyecto 46270.

Índice general

1. Introducción	1
1.1. Sobre la percepción humana del color	1
1.2. Segmentación	5
1.2.1. Segmentación en colores	6
1.3. La tesis y sus contribuciones	7
1.3.1. Motivación del tema de tesis	7
1.3.2. Objetivos de la tesis	8
1.3.3. Contribuciones de la tesis	8
2. Estado del arte de la temática	10
2.1. Espacios de colores	10
2.1.1. Espacio OSA	12
2.1.2. Espacio Munsell	12
2.1.3. Espacio RGB	13
2.1.4. Espacio HSI	14
2.1.5. Espacios CIE	16
2.2. Métodos de segmentación en colores. Enfoque clásico	18
2.3. Modelos de categorización del color	24
2.4. La segmentación en colores y los modelos de categorización del color	39
3. Sobre la propuesta de tesis	41

3.1. El enfoque Bayesiano junto a técnicas de interacción espacial	42
3.2. Campos aleatorios de medida Markovianos Gaussianos con entropía controlada .	45
3.3. Generalización de Campos aleatorios de medida Markovianos Gaussianos con entropía controlada	47
3.3.1. Estimación de la verosimilitud	50
3.4. Método de segmentación propuesto	55
3.4.1. Segmentación lingüística del color a través de un modelo Bayesiano jerárquico	56
4. Trabajo realizado y discusión	60
4.1. Construcción del conjunto de entrenamiento	60
4.2. Modelación de las categorías de color en un espacio a través de la estimación de la verosimilitud	63
4.2.1. Modelación en el espacio RGB	65
4.2.2. Modelación en los espacios $L^*a^*b^*$ y $L^*u^*v^*$	69
4.3. Segmentación lingüística de imágenes en colores mediante un modelo Bayesiano jerárquico	80
4.3.1. Uso directo del diccionario probabilístico	80
4.3.2. Estimación del atributo C	84
4.3.3. Estimación del atributo S	89
4.3.4. Comparación de la propuesta con otros métodos	91
5. Otras Aplicaciones	98
5.1. Filtrado de ruido en imágenes en colores	98
5.2. Compresión de imágenes	110
6. Conclusiones	113
A. Sobre splines	115
B. Experimentos	117

Índice de figuras

1.1. Espectro electromagnético.	2
1.2. Datos obtenidos por De Valois y colaboradores. En el eje X se tienen las longitudes de onda de los estímulos de luz usados, en el eje Y: la respuesta recibida ante cada estímulo. En cada gráfico aparecen 3 curvas para indicar la respuesta ante diferentes niveles de intensidad del estímulo luminoso.	3
1.3. Representación de las zonas del cerebro que intervienen en el procesamiento del color.	4
2.1. Representación cilíndrica del sistema Munsell.	13
2.2. Representación geométrica del RGB.	14
2.3. Representación geométrica del HSI.	15
2.4. (a) Representación gráfica del espacio $L^*u^*v^*$, (b) Representación gráfica del espacio $L^*a^*b^*$	18
2.5. Métodos de segmentación en colores comúnmente usados.	19
2.6. (a) Muestras de color usadas por Berlin y Kay en su experimento, (b) Categorización obtenida a partir de observadores de habla inglés, (c) Los puntos negros señalan los focos en la categoría.	26
2.7. Partición del espacio OSA, se muestra el plano de cromaticidad. Los círculos representan las zonas de consenso obtenidas. El nivel de intensidad es identificado por el diámetro de los círculos. Los discos negros representan la localización de los colores focales y los cuadrados la localización de los centroides.	27
2.8. Representación esquemática de los enlaces entre colores. El tamaño de las esferas representa la localización en el plano de intensidad. Blanco y negro no se consideran.	28

2.9. Zonas de consenso obtenidas por Sturges y Whitfield en el espacio de color de Munsell.	30
2.10. Respuestas espectrales, como funciones de membresía a las categorías roja, verde, azul y amarilla.	31
2.11. Los contornos (polígonos) y los focos (puntos) de las categorías básicas detectadas por Berlin y Kay, representadas en los espacios XYZ, $L^*a^*b^*$ y NPP (de izquierda a derecha).	33
3.1. (a) Imagen original, (b) segmentación usando ecuación (3.18), (c) segmentación usando ecuación (3.24)	49
3.2. Interpolación por GMMRF, a) $\lambda = 0.05$ b) $\lambda = 10$. En ambos casos el número de iteraciones es 2000.	52
3.3. Representación gráfica de la propuesta de segmentación.	59
4.1. Experimento de nombramiento de color considerando contexto.	61
4.2. Experimento de nombramiento de color sin contexto.	62
4.3. (a) Ubicación de una muestra de color en el espacio RGB, (b) respuesta de un observador en el experimento con contexto, (c) respuesta del observador en el experimento sin contexto.	62
4.4. Distribución de las muestras en los espacios (a) RGB, (b) $L^*a^*b^*$ y (c) $L^*u^*v^*$	64
4.5. Distribución de los nodos para los splines cuadráticos en el espacio RGB usando $\Delta = 64$	65
4.6. Resultados de la interpolación en el espacio RGB para la categoría azul usando $\tau = 0,01$ con (a) $\Delta = 16$, (b) $\Delta = 32$ y (c) $\Delta = 64$	66
4.7. Resultados de la interpolación por splines cuadráticos usando $\tau = 0$	66
4.8. Zonas de color asignadas a las categorías: (a) verde y (b) anaranjado ; (c) zona asignada a la categoría café.	67
4.9. (a) Partición del cubo RGB en 11 clases. Localización de las categorías (b) rojo, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) café y (j) negro , gris y blanco en el cubo RGB.	68

4.10. (a) Plano $B = 0$ en el espacio RGB, (b) partición obtenida, (c) representación de la entropía, (d) representación de la componente de mayor valor en el vector de verosimilitudes.	69
4.11. (a) Distribución de los nodos en el espacio $L^*u^*v^*$, (b) distribución de los nodos en el espacio $L^*a^*b^*$. Reticulado regular.	70
4.12. (a) Partición del espacio $L^*u^*v^*$ en 11 clases usando retícula regular. Localización de las categorías (b) roja, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) cafe y (j) negro, gris y blanco.	71
4.13. (a) Partición del espacio $L^*a^*b^*$ en 11 clases usando retícula regular. Localización de las categorías (b) roja, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) cafe y (j) negro, gris y blanco.	72
4.14. Otros experimentos en el espacio $L^*u^*v^*$, (a) $\tau = 0.01$, retícula de $5 \times 5 \times 10$, (b) $\tau = 0.1$, retícula de $5 \times 5 \times 10$, (c) $\tau = 1$, retícula de $5 \times 5 \times 10$, (d) $\tau = 0.01$, retícula de $10 \times 10 \times 10$, (e) $\tau = 1$, retícula de $10 \times 10 \times 10$	73
4.15. Otros experimentos en el espacio $L^*a^*b^*$, (a) $\tau = 0.01$, retícula de $5 \times 5 \times 10$, (b) $\tau = 0.1$, retícula de $5 \times 5 \times 10$, (c) $\tau = 0.01$, retícula de $10 \times 10 \times 10$, (d) $\tau = 1$, retícula de $10 \times 10 \times 10$	74
4.16. (a) Distribución de los nodos en el espacio $L^*u^*v^*$, (b) Distribución de los nodos en el espacio $L^*a^*b^*$. Los nodos son resultado de llevar las coordenadas de los nodos RGB a coordenadas $L^*u^*v^*$ y $L^*a^*b^*$	75
4.17. (a) Partición del espacio $L^*u^*v^*$ en 11 clases usando transformación de coordenadas de nodos RGB a nodos $L^*u^*v^*$. Localización de las categorías (b) roja, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) cafe y (j) negro, gris y blanco.	76
4.18. (a) Partición del espacio $L^*a^*b^*$ en 11 clases usando transformación de coordenadas de nodos RGB a nodos $L^*a^*b^*$. Localización de las categorías (b) roja, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) cafe y (j) negro, gris y blanco.	77

4.19. (a) Plano $B=0$ en el espacio RGB, (b) partición del espacio $L^*u^*v^*$ con retícula regular, y representada en el espacio RGB (plano $B=0$), (c) mapa de entropía, (d) componente de mayor valor en el vector de probabilidades, (e) partición del espacio $L^*u^*v^*$, usando reticulado no regular, y representada en el espacio RGB (plano $B=0$), (f) mapa de entropía, (g) componente de mayor valor en el vector de probabilidades.	78
4.20. (a) Plano $B=0$ en el espacio RGB, (b) partición del espacio $L^*u^*v^*$ (reticulado no regular), representada en el espacio RGB (plano $B=0$), (c) mapa de entropía, (d) componente de mayor valor en el vector de probabilidades, (e) partición del espacio $L^*a^*b^*$ (reticulado no regular), representada en el espacio RGB (plano $B=0$), (f) mapa de entropía, (g) componente de mayor valor en el vector de probabilidades.	79
4.21. (a) Imagen original. Segmentación usando la partición en b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$	81
4.22. (a) Imagen original, segmentación usando la partición en (b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$	81
4.23. (a) Imagen original, segmentación usando la partición en b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$	82
4.24. (a) Imagen original, segmentación usando la partición en (b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$	83
4.25. (a) Imagen original, segmentación usando la partición en (b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$	83
4.26. (a) Imagen original, (b) segmentación usando la partición en el espacio $L^*u^*v^*$	84
4.27. (a) Imagen original, (b) segmentación obtenida al estimar C a través de (3.17), (c) segmentación obtenida al estimar C usando (3.24).	85
4.28. (a) Zona del follaje ampliada en imagen original de la Figura 4.28, (b) ampliación sin considerar interacciones (expresión (3.17)), (c) ampliación considerando interacciones (expresión (3.24)).	86
4.29. (a) Imagen original, (b) segmentación obtenida al estimar C usando (3.24) sin considerar incertidumbre en zonas oscuras, (c) segmentación obtenida al estimar C usando (3.24) considerando incertidumbre en zonas oscuras.	87

4.30. (a) Imagen original, (b) segmentación obtenida al estimar C usando (3.24), sin considerar incertidumbre en zonas oscuras de color, (c) segmentación obtenida al estimar C usando (3.24), considerando incertidumbre.	88
4.31. (a) Imagen original, (b) segmentación obtenida al estimar C usando (3.24), (c) segmentación obtenida al estimar S dada la estimación de C , (d) representación de la estimación de C usando la información de color en la imagen, (e) representación del atributo S usando la información de color en la imagen. . . .	89
4.32. (a) Imagen original, (b) representación de la estimación de S dado C usando la información de color de imagen, (c) representación de la estimación de S dado C usando la información de color de imagen, luego de eliminar componentes conexas con área menor que cierto número de píxeles.	90
4.33. (a) Imagen original, (b) segmentación usando la nueva propuesta (solo se muestra representación del atributo C , usando expresión (4.3)), (c) bordes detectados a partir de la estimación del atributo C , (d) bordes detectados por Canny.	91
4.34. (a) Imagen original, (b) bordes obtenidos usando segmentación por <i>mean shift</i> , (c) bordes obtenidos usando algoritmo de gradiente de textura, (d) bordes obtenidos usando algoritmo basado en gradiente de color, (e) bordes obtenidos por la nueva propuesta.	93
4.35. (a) Imagen original, (b) bordes obtenidos usando segmentación por <i>mean shift</i> , (c) bordes obtenidos usando algoritmo de gradiente de textura, (d) bordes obtenidos usando algoritmo basado en gradiente de color, (e) bordes obtenidos por la nueva propuesta.	94
4.36. (a) Imagen original, (b) segmentación usando <i>mean shift</i> , (c) segmentación usando la nueva propuesta con representación a través de (4.3) considerando componentes conexas.	95
4.37. (a) Imagen original, (b) bordes detectados por la nueva propuesta, (c) segmentación manual dada en la base de datos de Berkeley.	96
4.38. (a) Imagen original, (b) bordes detectados por la nueva propuesta, (c) segmentación manual dada en la base de datos de Berkeley.	97

5.1.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 8$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	99
5.2.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 10$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	100
5.3.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 15$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	101
5.4.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 8$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	102
5.5.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 10$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	103

5.6.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 15$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	104
5.7.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 8$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	105
5.8.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 10$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	106
5.9.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 15$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	107
5.10.	(a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 15$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).	109

5.11. (a) Imagen original, (b) representación del atributo C. Refinamiento del atributo C a través de (c) 2 subclases, (d) 4 subclases y (e) 6 subclases usando la ecuación (4.3) en cada subclase encontrada. Refinamiento del atributo C a través de (f) 2 subclases, (g) 4 subclases y (i) 6 subclases usando la ecuación (4.3) en cada componente conexa.	111
5.12. (a) Imagen original, (b) representación del atributo C. Refinamiento del atributo C a través de (c) 2 subclases, (d) 4 subclases y (e) 6 subclases usando la ecuación (4.3) en cada subclase encontrada. Refinamiento del atributo C a través de (f) 2 subclases, (g) 4 subclases y (i) 6 subclases usando la ecuación (4.3) en cada componente conexa.	112
A.1. Ejemplos de uso de splines dados a) dos puntos, b) tres puntos, c) 4 puntos, d) n puntos	116
B.1. Leyenda empleada para la codificación de los atributos C y S durante la segmentación.	118
B.2. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	119
B.3. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	120
B.4. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	121
B.5. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	122
B.6. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	123

B.7. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	124
B.8. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	125
B.9. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	126
B.10.(a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	127
B.11.(a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	128
B.12.(a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	129
B.13.(a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).	130

Índice de Tablas

4.1. Por ciento que ocupa cada categoría en los espacios de color RGB, $L^*u^*v^*$ y $L^*a^*b^*$	80
--	----

Capítulo 1

Introducción

1.1. Sobre la percepción humana del color

El color es un fenómeno perceptual relacionado con la respuesta del ojo humano a diferentes longitudes de onda del espectro visible (ver Figura 1.1).

Acorde a las teorías postuladas, la visión del color inicia en la retina donde células fotoreceptoras llamadas conos, reciben los estímulos de luz. Lo dicho anteriormente se conoce como la teoría de la visión del color definida por Young-Helmholtz [1]. Estos autores establecieron que uno de los receptores es más sensible a la luz azul, otro a la verde y el tercero a la roja. A partir de aquí cualquier color puede ser descrito como una combinación de estas tres sensaciones.

Según lo postulado por Young-Helmholtz el color se considera como una propiedad puntual totalmente determinada por la luz incidente en un punto y luego interpretada en la retina. Sin embargo, esto contradice el fenómeno de contraste, que explica como el color percibido en un punto está determinado no sólo por la cualidad de la luz incidente, sino también por el contexto. Motivado por este último planteamiento, Hering [2] postula la existencia de tres pares de variables [3]: Y-B (amarillo-azul), R-G (rojo-verde), W-Bk (blanco-negro). Estas variables representaban las respuestas de tres procesos de oposición, que según Hering tenían lugar durante la percepción. Esta teoría fue rechazada por muchos investigadores porque en ella se postulaba la existencia de receptores químicos en la retina, que eran afectados en sentido opuesto por diferentes longitudes de onda de luz y no había evidencias experimentales sobre estos procesos en la etapa de recepción de un estímulo de luz.

En 1966, a través de experimentos neurofisiológicos, se obtuvieron evidencias de los procesos de oposición expresados por Hering después del nivel de recepción. De Valois y colaboradores

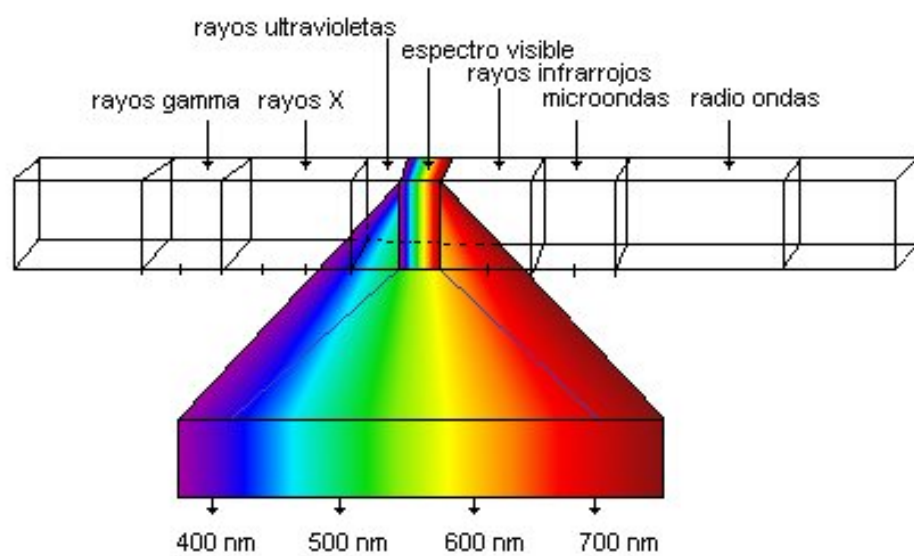


Figura 1.1. Espectro electromagnético.

[4], realizando estudios en cerebros de monos, descubrieron que en el LGN ¹, existen seis tipos de células, relacionadas con el procesamiento del estímulo de luz. Las respuestas espectrales de estas células aparecen en la Figura 1.2.

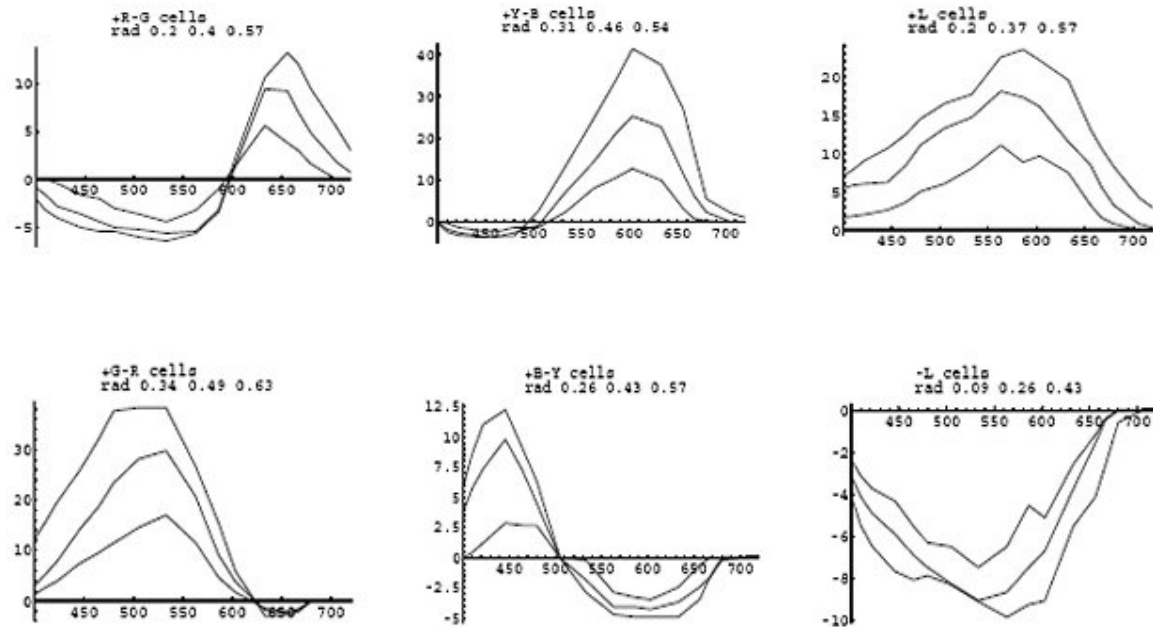


Figura 1.2. Datos obtenidos por De Valois y colaboradores. En el eje X se tienen las longitudes de onda de los estímulos de luz usados, en el eje Y: la respuesta recibida ante cada estímulo. En cada gráfico aparecen 3 curvas para indicar la respuesta ante diferentes niveles de intensidad del estímulo luminoso.

Luego de analizar los gráficos dados en la Figura 1.2, en [4] llegaron a la conclusión siguiente: entre los 6 tipos de células encontradas, 4 tenían una respuesta espectral de tipo “oponente o adversario” (vea los dos primeros gráficos de la primera fila y los dos de la segunda en la Figura 1.2). A estas células las llamaron +R-G, +G-R, +Y-B, +B-Y y las relacionaron con la información de color, en tanto que las otras dos estaban relacionadas con el procesamiento de la intensidad y estaban presentes en todo el espectro.

¹Las fibras del nervio óptico terminan en dos partes en un lugar del cerebro conocido como tálamo; estas partes se conocen como *Lateral Geniculate Nuclei*.

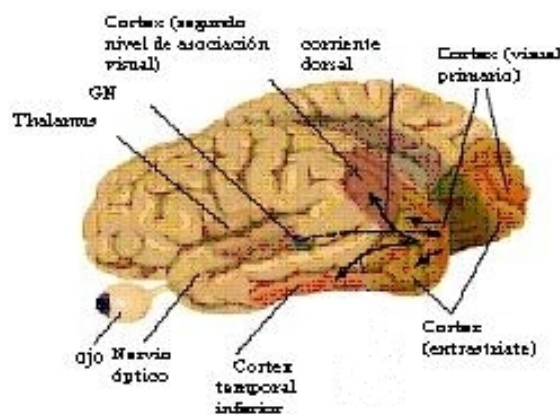


Figura 1.3. Representación de las zonas del cerebro que intervienen en el procesamiento del color.

Las investigaciones hechas en [4] permitieron conocer que después del nivel de recepción del estímulo de luz, las respuestas de los conos son codificadas en dos tipos de señales: acromática y cromática.

Luego de este proceso las señales obtenidas son interpretadas en el cerebro, específicamente en una zona conocida como corteza², considerando el contexto en que fue recibido el estímulo y la experiencia visual acumulada. Lo dicho anteriormente se resume en la Figura 1.3.

El proceso de percepción humana conduce a la categorización a través de etiquetas lingüísticas (palabras); el color no es una excepción [5], de ahí que la interpretación del estímulo luminoso se exprese a través de nombres o categorías de color. Aunque existen muchas representaciones

²En el neuroencéfalo se llama ‘corteza’(o ‘corteza’) al área más superficial del cerebro constituida por estratos o “capas”de redes neuronales, las que se observan como materia gris. El corteza se encuentra particularmente desarrollado en las aves y sobre todo en los mamíferos, siendo tal área histológica la responsable de las actividades psíquicas más elaboradas.

del color (espacios), en nuestra vida diaria identificamos los colores por sus nombres [5] . En 1969 Berlin y Kay [6], a través del estudio experimental de 20 lenguajes y de la investigación en la literatura relacionada con los términos de color de otros 78, presentaron evidencias para indicar que todos los lenguajes comparten un sistema universal de asignación de colores y que existen categorías básicas universales del color. Más tarde en 1972 [7] otros trabajos en esta área demostraron que las categorías universales son inherentes de la percepción humana porque la visión del color es el resultado de un proceso neurofisiológico común a todos los seres humanos.

1.2. Segmentación

Clásicamente, la segmentación se define como la partición de la imagen en sus partes constituyentes, siendo cada una de ellas homogéneas respecto a cierto atributo como por ejemplo: textura, intensidad y color [8]. Formalmente la definición es la siguiente [9]:

Si F es un conjunto de píxeles y $P()$ es un atributo de homogeneidad (uniformidad) para grupos de píxeles conectados , entonces la segmentación es la división del conjunto F en subconjuntos o regiones S_1, S_2, \dots, S_n tales que:

$$\bigcup_k S_k = F, k = 1, \dots, n, \quad (1.1)$$

$$S_i \cap S_j = \emptyset, i \neq j. \quad (1.2)$$

El atributo de uniformidad $P(S_i)$ se cumple para todos los elementos de las regiones S_i . Por otra parte $P(S_i \cup S_j)$ no se cumple cuando S_i y S_j son adyacentes. Siguiendo esta definición, en la literatura se registran diversos métodos y sus aplicaciones [9],[10],[11],[12], [13],[14],[15],[16], [17],[18],[19], [20],[21],[22]; obteniéndose como resultado la extracción de regiones en la imagen que resultan de interés según la tarea de estudio.

La segmentación es básicamente resultado de la percepción. Según este punto de vista la definición podría ser la siguiente [23]: *la segmentación es una transformación de colecciones de píxeles de la imagen en regiones y objetos visualmente significativos*. En el trabajo reportado en [23] el proceso de segmentación considera este aspecto.

Tanto el enfoque clásico, como el perceptual en el proceso de la segmentación son válidos e idealmente ambos deben dar salidas que se correspondan con el juicio del observador.

1.2.1. Segmentación en colores

La segmentación se realiza tanto para imágenes monocromáticas como en colores. El empleo del color representa un poderoso descriptor que a menudo simplifica la identificación de un objeto y su extracción de una imagen. Esto está en correspondencia con el hecho de que el ojo humano puede reconocer solamente dos docenas de niveles de gris; sin embargo puede diferenciar miles de matices e intensidades de color [24].

Básicamente la segmentación en colores parte de la segmentación en grises aplicada sobre diferentes espacios de color. Entre las técnicas creadas para el caso monocromático se encuentran: umbralado de histogramas, agrupamiento, crecimiento de regiones, detección de bordes, técnicas difusas (*fuzzy*), campos aleatorios Markovianos (CAM) junto a la estimación Bayesiana y redes neuronales [9],[12],[13],[15],[16],[10],[17],[18],[25].

La selección de un espacio o modelo de color es un aspecto crucial en el procesamiento en colores. En esencia un modelo de color es una especificación de un sistema tridimensional, en el que cada color queda definido por un único punto [24]. Una de las representaciones más comunes es la RGB, a partir de la cual se obtienen otras realizando transformaciones lineales y no lineales. Aún no se conoce un modelo de color que se considere como el más apropiado.

Una vía de extensión de los métodos de segmentación en grises, es aplicarlos sobre cada componente de una representación de color seleccionada. Posteriormente las salidas son combinadas de alguna manera para llegar al resultado final. Este proceder da lugar a un problema: *¿ cómo procesar la información de color como un todo para un pixel sin distorsionar la percepción del color por el observador humano ?*. Otra vía es la propuesta en [22] donde se procesa toda la información de color de un pixel.

Todavía no existe una teoría universal sobre la segmentación en colores; esto significa que aún no existen algoritmos generales, ni un espacio de color que sea adecuado para todo tipo de aplicación [8].

La segmentación tiene carácter perceptual, por eso no es posible darle una solución analítica única. Esto explica porque los enfoques difuso y probabilístico hayan sido exitosamente usados pues modelan la actividad cognitiva del ser humano. Con respecto a los espacios de colores, aquellos que describen mejor el proceso de percepción permiten resultados muy satisfactorios. Como se expresó en la sección 1.1 la percepción humana conduce a la categorización, lo que se traduce en el uso de etiquetas lingüísticas. Partiendo de los trabajos desarrollados por Berlin y Kay [6] se han desarrollado diferentes modelos de categorización del color [26], [27],[28],[5], [29] con los cuales se tiene una descripción lingüística del proceso de asignación de colores por el observador humano. Dichos modelos usan colores que se consideran como salientes

perceptuales para la descripción de cualquier otro color. Sería muy interesante, bajo el enfoque perceptual de la segmentación, la fusión de ésta con los modelos de categorización para dar lugar a una segmentación lingüística. Esto facilitaría la obtención de descriptores de objetos, generar anotaciones sobre la composición en color de una imagen y finalmente lograr una segmentación cuyos resultados se puedan interpretar en lenguaje natural usando un conjunto mínimo de palabras.

La segmentación es un problema mal planteado, es por eso que el uso de la teoría de campos ocultos Markovianos junto a la estimación Bayesiana ha resultado muy exitoso.

La combinación de la segmentación con los modelos de categorización del color y el uso de técnicas Bayesianas es precisamente el tema de estudio de este trabajo.

1.3. La tesis y sus contribuciones

1.3.1. Motivación del tema de tesis

Son varios los modelos de categorización del color que se proponen en la literatura para entender las etiquetas lingüísticas usadas por el hombre al asignar el color, sin embargo los mismos no son aplicados para la segmentación (vea sección 2). En algunos casos, aunque se propone su uso, la implementación se hace muy difícil por considerar parámetros físicos, cuya medición en condiciones normales se hace compleja. Por otra parte, la construcción de los modelos debe resolver los problemas siguientes:

- ¿ Cómo particionar un espacio de color en categorías básicas ?.
- ¿ Cómo expandir el concepto de colores básicos a un vocabulario más general, adecuado para diferentes aplicaciones, teniendo en cuenta no sólo a estos, sino modificadores de intensidad o de saturación tales como claro y pálido por ejemplo ?[5].
- ¿ Cómo expresar el grado de pertenencia a un color, de forma similar al dado por el observador humano? [5]. Este aspecto se hace aún más difícil en regiones de color muy oscuras.

Los modelos de categorización existentes, como se planteó anteriormente, no se usan directamente para la segmentación , sino como un medio de nombramiento del color de las regiones extraídas después de la segmentación por algún método. El carácter perceptual de la segmentación, y específicamente el caso en el que el color es el atributo de clasificación, hace

posible su fusión con un modelo que explique el proceso humano de nombramiento del color. Por otra parte si el método creado usa un enfoque Bayesiano, no sólo se consideran las observaciones en un pixel, sino también información de contexto de ese sitio. Además de este aspecto, se tiene que la solución obtenida bajo dicho enfoque considera la incertidumbre presente en el proceso de clasificación humana. Estas dos características describen elementos de la percepción, lo que explica el buen desempeño de las técnicas Bayesianas en visión computacional.

Todo lo planteado constituye la motivación de este trabajo de tesis titulado *Segmentación lingüística del color usando un modelo Bayesiano jerárquico de categorización*. Como resultado de esta propuesta se modela cada categoría o término básico de color en un espacio seleccionado. La modelación conlleva a un diccionario probabilístico de términos básicos de color que es usado para la segmentación usando técnicas Bayesianas.

1.3.2. Objetivos de la tesis

El objetivo de esta tesis es:

Realizar la segmentación lingüística de imágenes en colores mediante un diccionario probabilístico obtenido a través de la modelación de categorías básicas de color.

Los pasos para el cumplimiento de este objetivo son los siguientes:

- A. Modelación de cada categoría de color en el espacio seleccionado, obteniéndose un diccionario probabilístico.
- B. Segmentación regularizada de la imagen en colores, dando lugar a regiones con nombres de color correspondientes a las categorías básicas seleccionadas en el diccionario. En este paso se modela la incertidumbre en zonas oscuras al asignar colores.
- C. Subdivisión de cada categoría de color encontrada usando los modificadores *claro* y *oscuro*. Los resultados de la segmentación deben corresponderse con la descripción de composición de color dada por un observador.

1.3.3. Contribuciones de la tesis

Las contribuciones científicas de esta tesis son las siguientes:

- Ubicación de regiones de color en un espacio. Se propone un modelo de categorización del color, a través del cual se da una aproximación de la geometría y localización espacial de colores definidos como básicos en un espacio de color seleccionado. El modelo

no necesita conocer prototipos de cada categoría, lo que en muchas ocasiones hace compleja la modelación. Cada punto en el espacio de color está definido por un vector cuyas componentes indican el grado de membresía a una de las categorías básicas. La composición de este vector es similar al descriptor obtenido por un observador humano.

- Uso del modelo de categorización en la segmentación. Se fusiona el modelo de categorización elaborado con técnicas Bayesianas de interacción espacial para lograr la segmentación. Se propone una generalización del trabajo dado en [30], que permite considerar las interacciones entre colores. En la propuesta de segmentación se simula la incertidumbre presente en zonas oscuras de color.

La estructura de este manuscrito es como sigue: el capítulo 2 está dedicado al estado del arte de la segmentación, el capítulo 3 describe herramientas y aspectos teóricos de la propuesta. Seguidamente, en el capítulo 4 se detallan los experimentos desarrollados y los resultados obtenidos así como una discusión de los mismos. En el capítulo 5 se describen otras aplicaciones del método elaborado y en el 6 se dan las conclusiones.

Capítulo 2

Estado del arte de la temática

2.1. Espacios de colores

Existe un número considerable de representaciones o espacios de colores. El considerar una de ellas como la mejor es hasta el momento imposible. Esto explica el porque la selección de un espacio de color es un tema de estudio. La estructura de estos espacios, como se podrá apreciar, es una consecuencia de todas las investigaciones realizadas (ver sección (1.1)) para entender el mecanismo de percepción del color. Antes de especificar algunas de las representaciones, que fueron vistas en la elaboración de este trabajo, es necesario tener en cuenta los siguientes conceptos [31]:

- Luminancia, intensidad: cantidad de energía que percibe un observador de una fuente de luz.
- Brillo: atributo de la sensación visual que es proporcional a la energía irradiada por una fuente y es una medida subjetiva con la que decimos este objeto se observa brillante u opaco (oscuro). Una observación importante es que en la literatura consultada los términos intensidad, brillo y luminancia suelen aparecer como un mismo concepto.
- Matiz o tono (Hue): atributo del color que distingue una fuente y se corresponde con la longitud de onda dominante. Mediante este atributo los objetos se denotan como rojos, verdes, amarillos, púrpura, etc.
- Saturación: es un atributo de la sensación visual que mide la pureza de un color e indica la cantidad de luz blanca que contiene un matiz. Mediante este atributo los objetos se dicen ser azul pálido, azul fuerte, etc. [8].

- Cromaticidad: es un atributo de la sensación visual acorde con el cual un área observada aparece más o menos coloreada. La cromaticidad está determinada por el tono y la saturación.

Un término usado en esta sección es: *espacio perceptualmente uniforme*. Esto refiere a aquellos espacios donde la distancia percibida entre colores es equivalente a la calculada usando la métrica euclidiana.

Las representaciones del color se sustentan en propiedades físicas, fisiológicas o psico-visuales. En [32] se da la siguiente clasificación:

- Espacios primarios: están basados en la teoría tricromática (sección (1.1)), la cual asume que cualquier color está determinado por la mezcla, en cantidades apropiadas, de tres colores primarios.
- Espacios de intensidad-cromaticidad: en ellos una componente representa la intensidad y las otras dos la cromaticidad.
- Espacios perceptuales: cuantifican la percepción humana del color usando intensidad (I), tono (H) y saturación (S).
- Espacios de ejes independientes: son el resultado de métodos estadísticos aplicados sobre cualquiera de los espacios antes mencionados para dar lugar a espacios donde las componentes están menos correlacionadas.

Los espacios de color de color también se clasifican como lineales y no lineales a partir de la transformación de la que son resultado.

En la temática de espacios de colores se considera también el tipo de iluminación. Entre los iluminantes establecidos por la Comisión Internacional de Iluminación (CIE), se encuentran los siguientes [33]:

- Iluminante A: se conoce como un blanco caliente que corresponde a una temperatura de color de 2854°K ; una lámpara de filamento de tungsteno radia luz de esta tonalidad de blanco.
- Iluminante B: se aproxima a la luz solar directa del mediodía, se puede obtener en laboratorios de óptica partiendo de una fuente de iluminante A y utilizando filtros especiales. Su temperatura es de aproximadamente 4800°K .

- Iluminante C: algunas veces se le denomina luz diurna y tiene la misma calidad que la luz difusa del cielo nublado. Su temperatura es de 6770^0K .
- Iluminante D65: simula la luz diurna, que puede ser considerada como una muestra de luz solar directa y de luz de cielo difusa. Se obtiene por la mezcla de luz de los tres fósforos de un tubo de rayos catódicos de televisión en color.

2.1.1. Espacio OSA

El sistema OSA, fue desarrollado por el Comité de Escalas Uniformes de color de la Sociedad de Óptica de América (OSA) entre 1947 y 1974. La representación de este sistema viene dada por una retícula en un romboedro. Cada color está rodeado de 12 vecinos perceptualmente equidistantes de él. La posición de un color queda definida por coordenadas en tres ejes que se interceptan formando ángulos rectos. Los ejes son L: intensidad, J (del francés "jaune")- eje amarillo-azul, G: eje verde-rojo.

El eje L tiene valores entre -7 y 5. Aquellos colores que yacen en el plano $L = 0$, poseen intensidad media. Los planos paralelos a $L=0$ y por encima se denotan $L= 1, 2, \dots$, mientras que los que están por debajo se denotan $L=-1, -2, \dots$ [31].

Los niveles de grises se caracterizan por tener $J = G = 0$. Valores positivos de J con $G = 0$, indican amarillos o café, en tanto que valores negativos de J indican azules. Valores positivos de G, con $J = 0$ indican verdes, mientras que valores negativos representan rojos-púrpuras [31]. Una representación bidimensional del sistema OSA podrá ser apreciada mas adelante en la Figura 2.7. El modelo de color OSA es un ejemplo de la categoría de tipo intensidad - cromaticidad.

2.1.2. Espacio Munsell

El sistema Munsell, creado en 1969, es una de las representaciones más antiguas del color y de amplio uso por pintores y todos los que en su trabajo requieren describir el color. Sus dimensiones son: matiz (H - hue), valor (V- value) y croma (C). En el sistema Munsell hay 5 colores que se consideran como principales: rojo (R), amarillo (Y), verde (G), azul (B) y púrpura (P). El círculo de matices está dividido (vea Figura 2.1) en 100 partes iguales. El valor (V) representa el nivel de luminancia percibida y está entre 0 y 10, finalmente la dimensión croma representa la fortaleza del color (saturación). Su límite inferior es 0, representando así el caso acromático. El valor superior depende de V y de H [8]. Un ejemplo de la representación Munsell es este : 5GY 8/6, indicando que se tiene un color amarillo verdoso en la escala 5 con

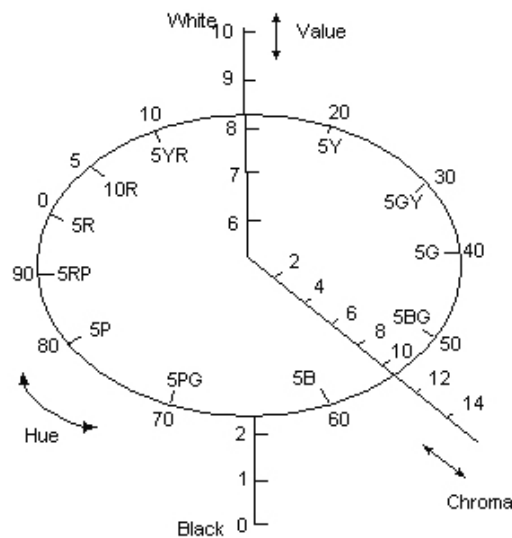


Figura 2.1. Representación cilíndrica del sistema Munsell.

valor 8 y croma 6. Se dice que este espacio es perceptualmente uniforme [31]. Un problema que presenta el sistema Munsell es la dificultad de encontrar las transformaciones a otros espacios. En este sentido en [34] se reporta un método para llevar imágenes en colores a dicho espacio. En la Figura 2.1 aparece una representación del sistema Munsell. Este sistema es otro ejemplo de la segunda categoría de tipos de representaciones de color.

2.1.3. Espacio RGB

En general, las imágenes en colores son adquiridas a través de la representación RGB debido a que dicho sistema es el más comúnmente usado en los sistemas de televisión y sistemas de adquisición de imágenes. El sistema RGB es un ejemplo de la primera categoría y su representación geométrica es la que aparece en la Figura 2.2. En el espacio RGB existe una alta dependencia entre sus componentes, lo que impide el procesamiento de un canal y omitir la información de otros. Esto no es absoluto, se registran aplicaciones donde si ha sido posible el procesamiento independiente bajo ciertas características de la imagen de estudio [35]. El sistema RGB no es perceptualmente uniforme y es muy sensible a cambios de intensidad. Una

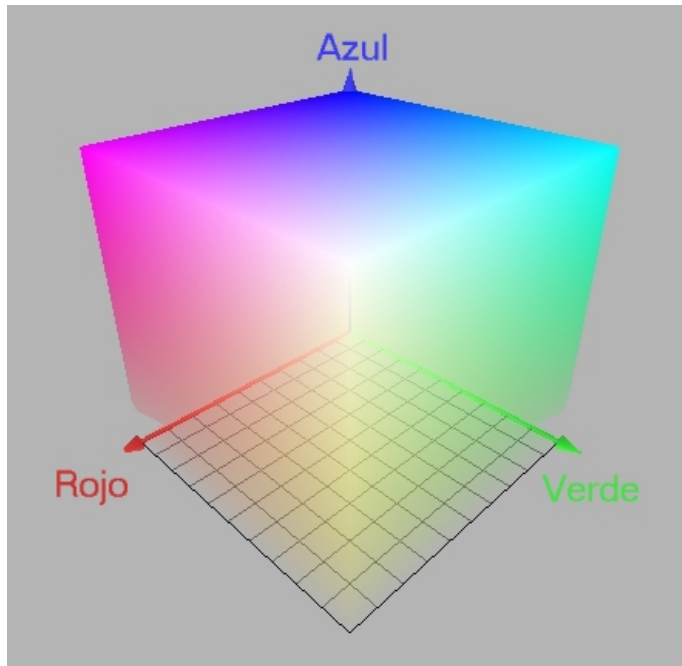


Figura 2.2. Representación geométrica del RGB.

vía para atenuar este problema es mediante la normalización de este espacio, como se indica a través de las siguientes expresiones:

$$r = \frac{R}{R + G + B}, \quad (2.1)$$

$$g = \frac{G}{R + G + B}, \quad (2.2)$$

$$b = \frac{B}{R + G + B}. \quad (2.3)$$

2.1.4. Espacio HSI

Existen variantes del sistema HSI: HSB (hue-saturation-brighthness), HSL (hue-saturation-lightness) y HSV(hue-saturation-value)[8]. En el HSI se separa la información de color de la de intensidad. La información de color está dada por las componentes H y S y la de intensidad por I. La descripción geométrica del sistema HSI según [24] aparece en la Figura 2.3. H es el ángulo del vector mostrado con respecto al rojo. Así cuando $H=0$ el punto P corresponde al color rojo. La saturación (S) es proporcional a la distancia desde P hasta el centro del triángulo e indica

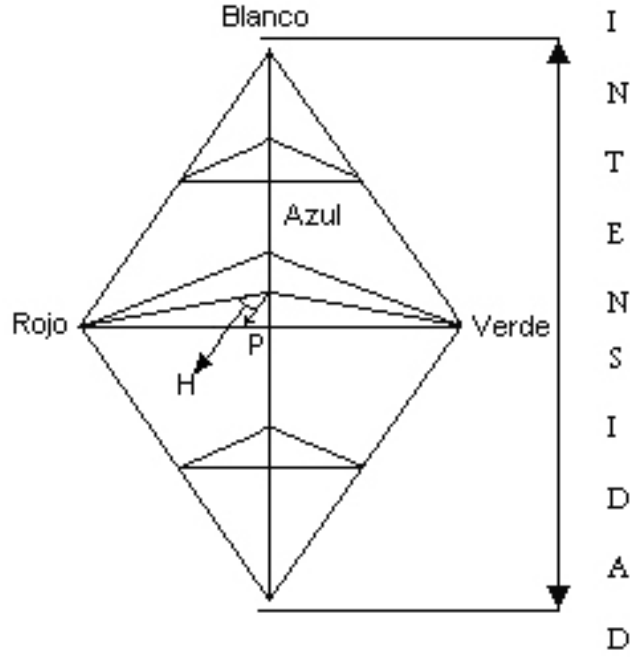


Figura 2.3. Representación geométrica del HSI.

la pureza del color. La intensidad, se mide con respecto a la línea perpendicular al plano del triángulo y que pasa por su centro. Las intensidades van desde el negro hasta llegar al blanco. El sistema de color HSI es una transformación no lineal del RGB [8].

$$H = \arctan\left(\frac{\sqrt[3]{3}(G - B)}{(R - G) + (R - B)}\right), \quad (2.4)$$

$$I = \frac{R + G + B}{3}, \quad (2.5)$$

$$S = 1 - \frac{\min(R, G, B)}{I}. \quad (2.6)$$

En el espacio HSI la información de color está separada de la de intensidad, por eso es posible realizar el procesamiento del color usando solamente las componentes de cromaticidad, o sea matiz y saturación. El matiz por ejemplo es muy útil, cuando hay cambios de iluminación para los que el mismo es invariante. En este caso se logra un procesamiento en 1D, lo que computacionalmente es menos costoso que en 3D como en el caso del sistema RGB. Una desventaja de esta representación está dada por las singularidades que aparecen en la representación cerca del eje de intensidades [8]. El HSI tiene la capacidad de representar los

colores de forma similar al hombre: distinguimos diferentes matices, y la percepción de diferentes niveles de intensidad y saturación no implica el reconocimiento de colores diferentes [24].

2.1.5. Espacios CIE

Los espacios de color desarrollados por la Comisión Internacional de Iluminación (CIE-Commission International de l'Éclairage) tienen la capacidad de representar las diferencias de color percibidas a través de la distancia euclidiana y son considerados como una aproximación del sistema visual humano [31]. En este espacio existen tres primarios (primarios artificiales) denotados como X, Y y Z [31], a partir de los cuales cualquier color se representa como una combinación de ellos. El espacio CIE XYZ es el resultado de mediciones directas sobre el ojo humano hechas a finales de 1920 por W. David Wright [36] y John Guild [37] y sirve de base para otras representaciones de color. Los valores de XYZ se calculan por medio de la siguiente transformación lineal del RGB.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (2.7)$$

En la expresión (2.7), los valores de R, G y B están en el intervalo [0,1]. Los elementos de la matriz de transformación dependen del tipo de iluminante empleado [33],[31] y estos valores están tabulados [31],[38]. A partir del XYZ, por una transformación no lineal, se calculan los valores en los espacios $L^*u^*v^*$ y $L^*a^*b^*$. La definición del $L^*a^*b^*$ es la siguiente:

$$L^* = 116 \left(\sqrt[3]{\frac{Y}{Y_0}} \right) - 16, \quad (2.8)$$

$$a^* = 500 \left(\sqrt[3]{\frac{X}{X_0}} - \sqrt[3]{\frac{Y}{Y_0}} \right), \quad (2.9)$$

$$b^* = 200 \left(\sqrt[3]{\frac{Y}{Y_0}} - \sqrt[3]{\frac{Z}{Z_0}} \right). \quad (2.10)$$

En las expresiones (2.8), (2.9) y (2.10) X_0 , Y_0 y Z_0 son los valores correspondientes al blanco en RGB.

La definición del espacio $L^*u^*v^*$ es:

$$L^* = 116 \left(\sqrt[3]{\frac{Y}{Y_0}} \right) - 16, \quad (2.11)$$

$$u^* = 13L^* (u' - u_0), \quad (2.12)$$

$$v^* = 13L^* (v' - v_0). \quad (2.13)$$

Los valores de u' y v' se calculan como:

$$u' = \frac{4X}{X + 15Y + 3Z}, \quad (2.14)$$

$$v' = \frac{9Y}{X + 15Y + 3Z}. \quad (2.15)$$

En (2.11) y (2.12) u_0 y v_0 se calculan según (2.14) y (2.15) respectivamente, usando los valores X_0 , Y_0 y Z_0 .

Tanto en el espacio $L^*a^*b^*$ como en el $L^*u^*v^*$ la componente L es la de luminancia. $L^* = 0$ indica negro y $L^* = 100$ indica blanco. Las componentes a^* y u^* representan la posición entre el magenta y el verde. Valores negativos de a^* y u^* indican cuanto hay de verde, mientras que valores positivos indican cuanto hay de magenta. Las componentes b^* y v^* representan la posición entre amarillo y azul. Valores negativos de b^* y v^* indican cuanto hay de azul, mientras que valores positivos indican cuanto hay de amarillo.

La relación no lineal en las expresiones que definen los espacios $L^*a^*b^*$ y $L^*u^*v^*$ tienen como objetivo simular la respuesta logarítmica del ojo humano. Se conoce experimentalmente que la respuesta R de un sistema sensorial no es proporcional a una entrada I , sino al logaritmo, o sea $R \propto \ln I$.

Las expresiones que definen los espacios $L^*a^*b^*$ y $L^*u^*v^*$ se corresponden con las investigaciones hechas por Stevens en 1975 [39], según el cual la respuesta ante un estímulo sigue la ley de la potencia, o sea:

$$\psi(x) = bx^p. \quad (2.16)$$

En (2.16) x es el estímulo recibido, b es una constante y $\psi(x)$ es la respuesta ante el estímulo. Según Stevens p depende del órgano sensorial; así para el oído $p = 0.67$, para el ojo $p = 0.33$ y para el olfato $p = 0.6$. Los valores de p varían entre individuos; pero estos valores dados son los más representativos. Además de la expresión (2.16) se sugieren otras tres que simulan el

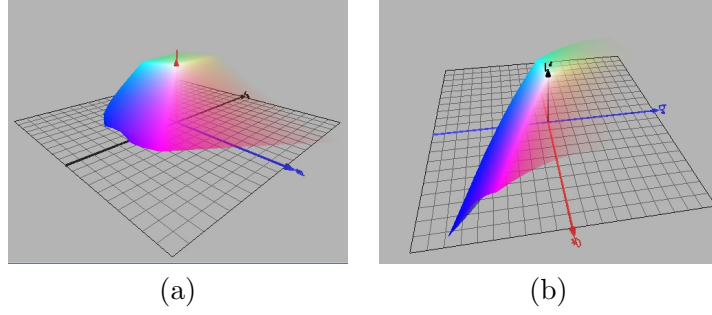


Figura 2.4. (a) Representación gráfica del espacio $L^*u^*v^*$, (b) Representación gráfica del espacio $L^*a^*b^*$.

proceso de percepción. Estas son:

$$\psi(x) = b(x - x_0)^p, \quad (2.17)$$

$$\psi(x) = bx^p - \psi_0, \quad (2.18)$$

$$\psi(x) = b[(x + x_0)^p - x_0^p]. \quad (2.19)$$

x_0 es el valor del estímulo cuando comienza la percepción y ψ_0 representa efecto de contraste. Comparando estas expresiones con las que determinan los espacios $L^*u^*v^*$ y $L^*a^*b^*$, se observa correspondencia, lo que explica el buen desempeño de estos espacios.

Los espacios $L^*a^*b^*$ y $L^*u^*v^*$ son perceptualmente uniformes. En la Figura 2.4 aparece la representación gráfica de los espacios comentados.

2.2. Métodos de segmentación en colores. Enfoque clásico

En la Figura 2.5 aparecen de forma resumida los métodos de segmentación aplicados en imágenes en colores siguiendo la definición clásica [8].

Generalmente los métodos de segmentación para imágenes monocromáticas se aplican directamente sobre cada componente del espacio de color seleccionado y luego los resultados se combinan de alguna manera, observándose como inconveniente que la información de color como un todo es omitida. Otro proceder es considerar al mismo tiempo todas las componentes de color, aumentándose así el costo computacional. A continuación, siguiendo el esquema en la Figura 2.5, se comentan algunas estrategias de segmentación en las que se ejemplifican ambos



Figura 2.5. Métodos de segmentación en colores comúnmente usados.

caminos.

Los métodos de umbralado asumen que las imágenes están compuestas por regiones de diferentes niveles de grises. Esto se refleja en histogramas formados por un cierto número de picos (modas) y valles, donde cada pico corresponde a una región. Se asume entonces que es posible encontrar un umbral ubicado en el valle entre dos picos que permite realizar agrupamiento (*clustering*) y detectar así regiones. En el caso de imágenes en colores se tiene un histograma multidimensional y, por tanto la división del espacio en varios grupos, implica una selección de picos significativos más laboriosa. Una vía para solucionar esta dificultad es operar sobre cada componente de color. En [17] construyen el histograma en cada componente del espacio RGB, luego hacen uso del espacio de escala para encontrar los picos significativos en cada uno; esta etapa la consideran como un paso grueso de la segmentación. Posteriormente los picos significativos se usan para refinar la segmentación haciendo uso de campos aleatorios Markovianos. En [40] se sugiere un histograma multidimensional usando tres espacios de color: RGB, YIQ y HSI. Un problema de la técnica de umbralado de histogramas es la selección de picos. Un criterio para la detección de picos es el dado en [41]:

$$S = \frac{T_p}{W_p}. \quad (2.20)$$

En la expresión (2.20) T_p indica el número total de píxeles entre dos valles y W_p indica el ancho del área del pico. Si S es mayor que cierto valor el pico se elige como significativo. El espacio utilizado en este trabajo fue el HSI. En estas propuestas mencionadas el umbralado se aplica sobre una sola componente. Otra solución es la de [42] y [18] al utilizar análisis de componentes principales en el espacio RGB y trabajar sólo en las componentes de mayor

importancia. Una propuesta interesante es la de [22] al elaborar una técnica no paramétrica para analizar un espacio de características complejo, como lo es por naturaleza, el espacio en que se representan las imágenes en colores. El algoritmo emplea la media desplazada (*mean shift*) [43] para el cálculo de modas en una distribución desconocida en el espacio de características $L^*u^*v^*$, considerando todas las componentes de dicho espacio al mismo tiempo e información espacial. La interpretación de la media desplazada es la siguiente: *mean shift* es un vector alineado respecto al gradiente local y define el camino hacia los puntos estacionarios de una función de densidad. Las modas de la función de densidad son los puntos estacionarios [22]. La distribución de los datos en el espacio de color, es tomada como la función de densidad a la cual se le detectan las modas; las mismas se corresponden con las regiones de color más significativas de la imagen.

En general, determinar el número de grupos (*clustering validity*) mediante un esquema no supervisado de segmentación sigue siendo un problema. En [44] tratan de resolver esto a través de promedios c-difusos (*fuzzy c-means*), luego de obtener una presegmentación por el umbralado en los histogramas de la imagen; *fuzzy c-means* identifica los grupos más significativos de la imagen.

Los métodos de umbralado basados en el histograma tienen la ventaja de no necesitar información a priori de la imagen y para un amplio rango de imágenes que cumplen los requerimientos de este método el mismo se desempeña muy bien. Una desventaja del método es que no considera la información espacial, por lo que en presencia de ruido son muy sensibles. Los métodos de crecimiento, división y fusión de regiones si consideran la información espacial y agrupan píxeles atendiendo a un criterio de similitud u homogeneidad, como por ejemplo: intensidad, textura o color. El problema de estas técnicas reside en la selección de un buen criterio de homogeneidad que distinga cada región. En [12] se sugiere uno considerando la varianza y la discontinuidad de intensidades:

$$H_{ij} = 1 - \left(\frac{S_{ij}}{S_{max}} \right) \left(\frac{D_{ij}}{D_{max}} \right). \quad (2.21)$$

En (2.21) S_{ij} es la desviación estándar calculada en una vecindad del píxel (i,j) y D_{ij} es una medida de la discontinuidad :

$$D_{ij} = \sqrt[2]{G_x^2 + G_y^2}. \quad (2.22)$$

G_x y G_y son las derivadas en las direcciones horizontales y verticales respectivamente. Los valores de H_{ij} están entre 0 y 1 y mientras más cercanos a 1 sean, la región alrededor del píxel

(i,j) es más homogénea ; S_{max} y D_{max} son los valores de desviación estándar y discontinuidad máximos respectivamente, en la vecindad del pixel (i,j) . En [12] se combina la técnica de umbralado con la técnica de crecimiento de regiones usando como representación del color el espacio $L^*a^*b^*$. Los métodos basados en crecimiento de regiones son muy dependientes de los puntos “semillas” o puntos de inicio de crecimiento, y del orden en que se recorran los píxeles en la imagen, lo que constituye una desventaja. Cuando se aplica el método de división en regiones, la semilla inicial es toda la imagen. Si la imagen no es homogénea, esta se divide en 4 regiones, siendo cada una de ellas una nueva semilla. El proceso de división termina cuando cada subregión obtenida sea homogénea. Una desventaja de este método es que las regiones obtenidas tienden a parecerse a la estructura de datos utilizada y este caso aparecen regiones en forma de cuadrados. Por lo general el método de división de regiones se combina con fusión para unir regiones similares. Las técnicas de crecimiento de regiones trabajan muy bien en imágenes con zonas homogéneas y tienden a ser menos sensible al ruido, porque los criterios de homogeneidad se determinan estadísticamente. Los métodos de crecimiento de regiones son mejores que los de umbralado o *clustering* porque consideran no sólo cierto espacio de características sino también información espacial [8].

Otra técnica usada es la basada en detección de bordes. Para imágenes en tonos de gris la detección de bordes está basada en encontrar las discontinuidades en los niveles de gris tratando así de localizar píxeles con cambios bruscos en el nivel de intensidad. Entre las técnicas de detección de bordes se encuentran las basadas en la elaboración de filtros pasa alto en el espacio de Fourier, teniendo en cuenta que las altas frecuencias espaciales están asociadas a cambios bruscos de intensidad. En este caso el problema radica en el diseño del filtro apropiado. Otros operadores usados son el de Robers, Sobel, Prewitt y el Laplaciano [24]. Los tres primeros se conocen como operadores de primeras diferencias y el último como operador de segundas diferencias. Todos son muy sensibles al ruido por considerar pequeñas ventanas y asumir que el borde se produce cuando hay cambios bruscos de intensidad. En el caso de imágenes en colores el borde se define como una discontinuidad en un espacio tridimensional. En [45] se elaboró un detector de bordes para color usando un operador de diferencias. Para cada pixel (x,y) se definen $HOE(x,y)$, $VOE(x,y)$, $NOE(x,y)$ y $SOE(x,y)$ las cuales determinan la fortaleza de los bordes en las direcciones horizontal, vertical, diagonal noreste y diagonal

norroeste respectivamente. El cálculo se realiza como se indica a continuación:

$$\begin{aligned} HOE(x, y) = & |I(x-1, y-1) + 2I(x, y-1) + I(x+1, y-1) - I(x-1, y+1) \\ & - 2I(x, y+1) - I(x+1, y+1)|, \end{aligned} \quad (2.23)$$

$$\begin{aligned} VOE(x, y) = & |I(x-1, y-1) + 2I(x-1, y) + I(x-1, y+1) - I(x+1, y-1) \\ & - 2I(x+1, y) - I(x+1, y+1)|, \end{aligned} \quad (2.24)$$

$$\begin{aligned} NOE(x, y) = & |2I(x-1, y-1) + I(x, y-1) + I(x-1, y) - I(x+1, y) \\ & - I(x, y+1) - 2I(x+1, y+1)|, \end{aligned} \quad (2.25)$$

$$\begin{aligned} SOE(x, y) = & |I(x, y-1) + 2I(x+1, y-1) + I(x+1, y) - 2I(x-1, y+1) \\ & - 2I(x-1, y) - I(x, y+1)|. \end{aligned} \quad (2.26)$$

En las expresiones anteriores I define el contenido de una componente de color en el espacio YUV, seleccionado como la representación de color en [45]. La fortaleza del borde se calcula como sigue:

$$MOE(x, y) = \max(HOE(x, y), VOE(x, y), NOE(x, y), SOE(x, y)). \quad (2.27)$$

Un pixel es borde si para cada canal $MOE(x, y) \geq Th$, donde Th es un umbral.

Existen trabajos en los que se demuestra que la cantidad de bordes detectados usando la componente de brillo en un espacio de color es mucho mayor que los obtenidos al usar las componentes de color, debido a que el color es relativamente invariante ante los cambios de iluminación [8]. Es por esta razón que en [46], usando la componente I en el espacio HSI, detectan en un primer paso bordes a través del operador de Canny [47] y luego se eliminan aquellos cuyas variaciones en contenido de H sean muy pequeñas. Los métodos basados en bordes son sólo la primera parte del proceso de segmentación. Ellos proporcionan una información muy útil sobre fronteras de objetos, la cual es usada por otras técnicas para finalmente lograr la segmentación. En [48], por ejemplo, se combina la información de borde con la de región. Las *semillas* para el crecimiento de regiones son generadas luego que se usa un detector de bordes en color; en el que los cambios en intensidad se pesan de forma diferente con respecto a los cambios en color. El espacio usado fue el RGB y la expresión que define el

detector de bordes es la siguiente:

$$dR_i(i, j) = R(i-2, j) + \frac{R(i-1, j)}{2} - \frac{R(i+1, j)}{2} - R(i+2, j), \quad (2.28)$$

$$dI_i = dR_i + dG_i + dB_i, \quad (2.29)$$

$$d_i = (dB_i - dG_i)^2 + (dR_i - dB_i)^2 + (dG_i - dR_i)^2,$$

$$d_j = (dB_j - dG_j)^2 + (dR_j - dB_j)^2 + (dG_j - dR_j)^2,$$

$$DC = \sqrt{d_i + d_j}, \quad (2.30)$$

$$dt = dI_i^2 + dI_j^2 + 3,0DC. \quad (2.31)$$

En (2.28) se expresa la medida de cambios en un canal en la dirección i , en (2.29) se define la medida para detectar cambios en intensidad; (2.30) define los cambios en color y finalmente (2.31) determina el detector de bordes creado. Una vez determinados los bordes, estos actúan como barreras durante el crecimiento de regiones.

Los métodos de segmentación basados en la detección de bordes trabajan bien para imágenes con un buen contraste, en caso contrario se hace muy difícil, por otro lado no siempre es posible obtener contornos cerrados. La ventaja de estos métodos está dada por el hecho de reflejar la forma en que los humanos percibimos los objetos: buscamos fronteras.

Otro camino usado para las segmentación es el uso de redes neuronales. Las mismas están formadas por un conjunto de elementos cuyos enlaces están determinados por pesos variables; así la red puede representar interacciones entre elementos y modelar cualquier tipo de función que describa tareas tales como la segmentación. Las redes neuronales son un método supervisado de clasificación, de ahí la etapa de entrenamiento, lo que las hace costosas; sin embargo luego de un buen entrenamiento la etapa de clasificación es muy rápida.

Entre los ejemplos de redes usadas están las *Hopfield* [49]. En [25] se localizan los picos significativos por el análisis de histograma; con esta información se tienen entonces las clases presentes. Posteriormente se construyen tres redes neuronales de tipo *Hopfield*, una por cada canal del espacio seleccionado. Los resultados en cada componente se combinan. Los espacios usados fueron RGB y la transformación de Karhunen -Loeve sobre el espacio RGB.

Los métodos de segmentación, cuya respuesta sobre la pertenencia de un pixel a una región es *dura*, no siempre reflejan correctamente lo que ocurre en una imagen real pues las regiones siempre no están bien definidas. Los métodos de segmentación basados en teoría de conjuntos difusos permiten considerar la ambigüedad e incertidumbre presentes en los problemas de reconocimiento. En este caso cada pixel en la imagen pertenece con cierto grado a una de las regiones. Ese grado se define por una función llamada función de membresía.

En [10] se presenta un método iterativo, en cada paso del cual se minimiza una función objetivo, con la cual se logra hacer mínima la distancia entre datos y los centros de clase, mientras que la distancia entre los centros de clase se hace máxima. En [50] se usa el concepto de conectividad difusa para la segmentación de varios objetos simultáneamente. La conectividad difusa determina la relación entre los píxeles de la imagen. Una vía para definirla podría ser el recíproco de la distancia entre ellos, haciendo una normalización entre 0 y 1. A partir de un número de *semillas* dado se calcula la conectividad difusa de cada píxel respecto a las *semillas*. Finalmente en cada objeto segmentado quedarán los píxeles cuya conectividad difusa respecto a las semillas asociadas a ese objeto sea mayor.

La segmentación, al igual que otras tareas en visión, puede ser formulada de forma natural como un problema de minimización de cierta función de energía [51]. Bajo este enfoque se busca el etiquetado f que minimice la función de energía dada por (2.32):

$$E(f) = E_{suave}(f) + E_{datos}(f). \quad (2.32)$$

En (2.32) f es suave y consistente con los datos observados; E_{suave} regula la suavidad en f , y E_{datos} controla la diferencia entre el etiquetado f y las observaciones. Geman y Geman [52] dieron una interpretación Bayesiana de las funciones de energía y propusieron un enfoque basado en campos aleatorios Markovianos. Basado en este trabajo se reportan otros [53], [54], [55], [30] en los que la combinación de técnicas Bayesianas con campos aleatorios Markovianos ha sido exitosamente usada para la segmentación. Los resultados de esta combinación indican la incertidumbre presente en la segmentación, y por tanto reflejan el carácter perceptual de la misma.

2.3. Modelos de categorización del color

Berlin y Kay al introducir el concepto de categorías básicas del color identificaron 11 términos básicos. De ellos 3 representan categorías acromáticas y 8 son cromáticas. Entre las características más importantes de los términos básicos de color se encuentran las siguientes [6]:

- Son monolémicos, es decir a través de una sola palabra expresan su significado.
- Su uso no está restringido a una clase de objetos (por ejemplo color ladrillo o color limón, en cambio si se dice anaranjado o verde se pueden describir muchos objetos incluyendo ladrillos y limones).

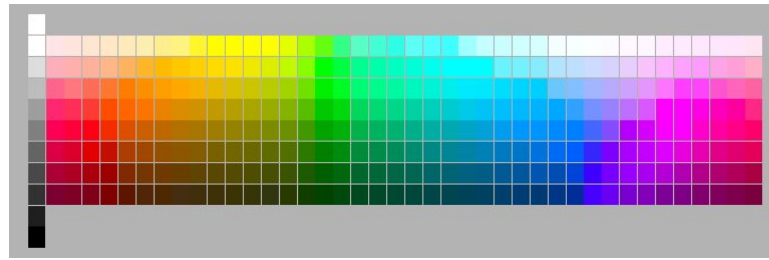
- Psicológicamente son salientes de información y puntos de referencia en la descripción de otro color.

En la primera parte del experimento de Berlin y Kay, los participantes definieron los nombres de los colores que eran básicos para ellos; seguidamente a todos se les mostró un total de 329 colores (vea Figura 2.6) usando el sistema Munsell (ver 2.1.2). Los observadores fueron instruidos para indicar el punto focal o prototipo (el mejor representante de cada categoría o color) y las fronteras de cada grupo de color encontrado (selección de otras muestras de color similares al punto focal). Como resultado de este experimento se hicieron las siguientes conclusiones:

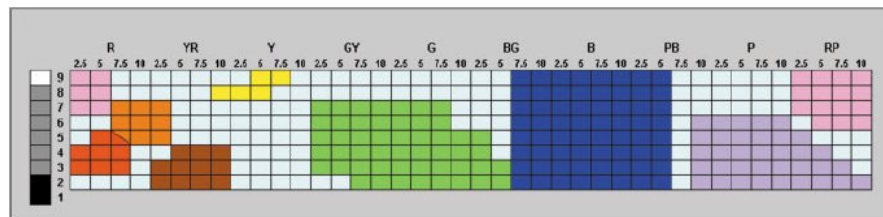
- Los puntos focales juegan un papel muy importante en la representación interna de cada color y la asignación a una u otra categoría se realiza con respecto a ellos.
- Las fronteras entre categorías varían entre lenguajes e individuos.
- Aunque lenguas diferentes usan diferente número de categorías básicas en su vocabulario, existe un total de 11 categorías universales: negro, blanco, gris, rojo, verde, amarillo, azul, café, rosado, anaranjado y púrpura.

Los cuadros de color blanco en las Figuras 2.6(b) y (c) corresponden a colores no asociados a ninguna categoría. Los resultados obtenidos por Berlin y Kay constituyeron el punto de partida para las investigaciones que se hacen en la temática categorización del color. Partiendo de ellos, en 1987 Boynton y Olson [26] confirmaron la diferencia entre términos básicos y términos no básicos del color (ejemplo: salmón, cyan, violeta) y definen la localización de los 11 colores básicos definidos por Berlin y Kay en el espacio OSA. En el experimento participaron 7 observadores y se usaron 424 muestras. Los resultados fueron los siguientes:

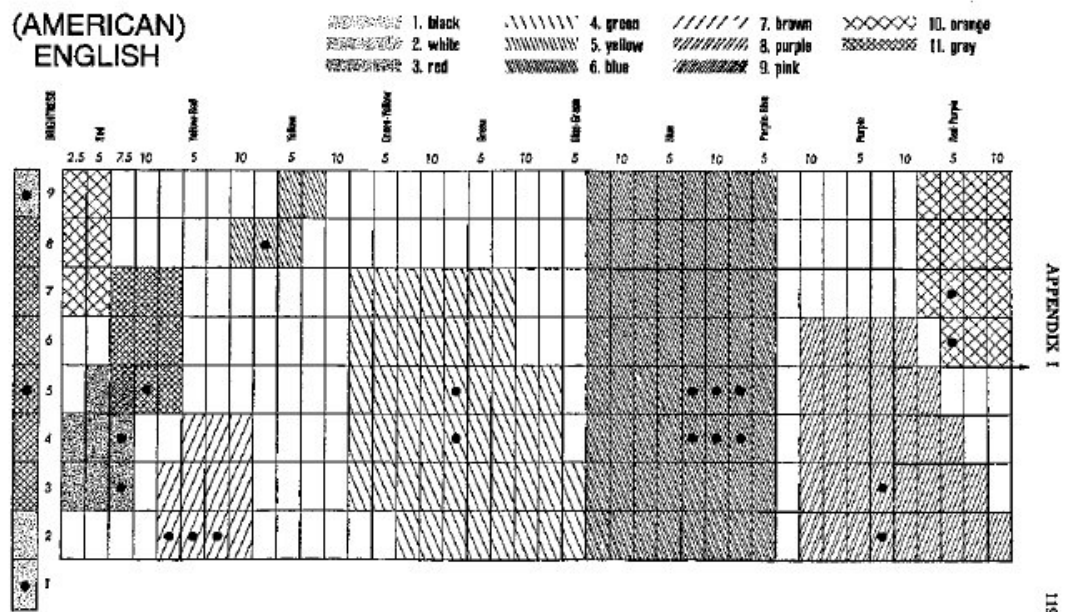
- Los colores ocupan zonas del espacio y no puntos aislados.
- Ubicación de las zonas de consenso en la clasificación de las muestras dentro del espacio OSA. Las zonas de consenso son aquellas correspondientes a muestras de color a las que todos los observadores le asignaron el mismo nombre. Estas áreas de consenso coincidieron con las asociadas a las 11 categorías básicas establecidas en [6].
- Los colores focales o prototipos están en las zonas de consenso. Se calculan los centroides, los cuales indican la tendencia central de cada categoría.
- El uso de términos no básicos hace más lenta la respuesta de los observadores.



(a)



(b)



(c)

Figura 2.6. (a) Muestras de color usadas por Berlin y Kay en su experimento, (b) Categorización obtenida a partir de observadores de habla inglés, (c) Los puntos negros señalan los focos en la categoría.

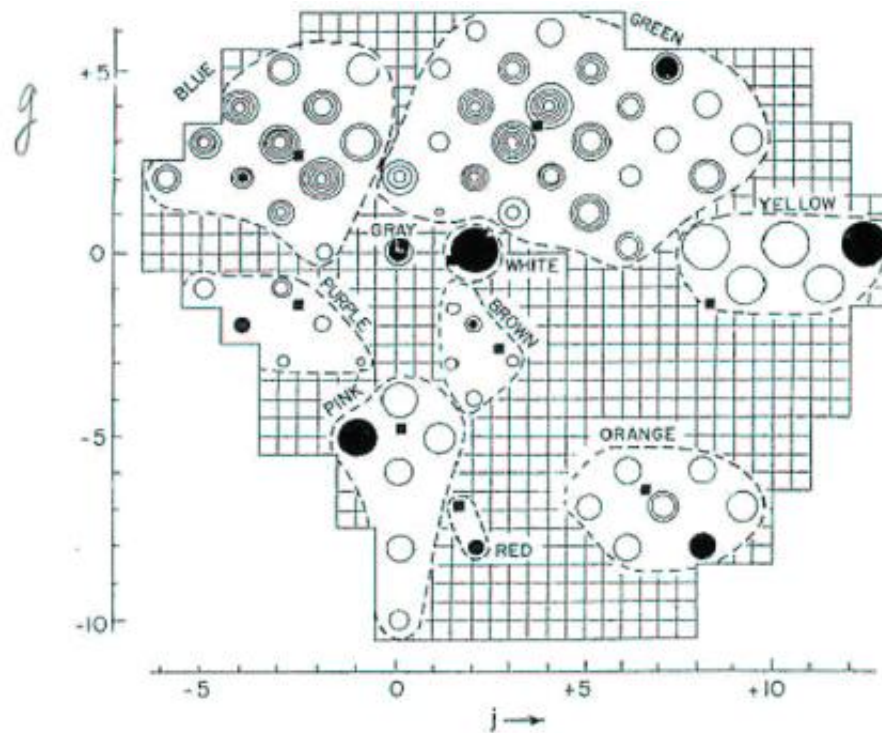


Figura 2.7. Partición del espacio OSA, se muestra el plano de cromaticidad. Los círculos representan las zonas de consenso obtenidas. El nivel de intensidad es identificado por el diámetro de los círculos. Los discos negros representan la localización de los colores focales y los cuadrados la localización de los centroides.

- Las áreas de consenso asociadas al verde y azul se presentan en casi todos los niveles de intensidad y ocupan la mayor parte en la partición del espacio. En cambio, la localización del resto de las categorías, está determinada tanto por el nivel de intensidad como por su posición en el plano cromático (ver figura 2.7).
- Entre algunas categorías existen enlaces, debido a que hay zonas donde la descripción del color se realiza usando más de una de ellas (ver Figura 2.8).
- La no existencia de enlaces entre categorías se justifica por su incompatibilidad perceptual, por ejemplo no hay enlaces entre el rojo y el verde.

En 1995 J. Sturges y Whitfield [29] realizaron una extensión de este trabajo, llevando a cabo los mismos experimentos pero en el espacio Munsell. La selección del espacio se justifica por el hecho de considerarse este sistema como destacado en estudios del lenguaje, además de que

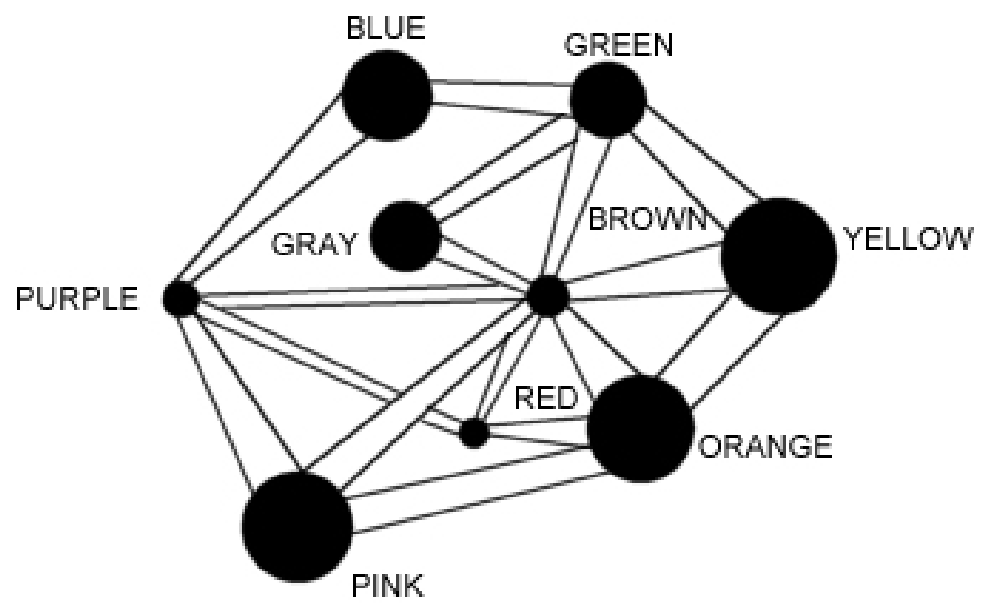


Figura 2.8. Representación esquemática de los enlaces entre colores. El tamaño de las esferas representa la localización en el plano de intensidad. Blanco y negro no se consideran.

es posible lograr en él una mejor representación de colores en diferentes niveles de intensidad y cromaticidad. En el trabajo de Boynton y Olson se planteó la necesidad de otra categoría para eliminar la dificultad del nombramiento del color en zonas entre amarillo, café, rosado, blanco y anaranjado; Sturges y Whitfield propusieron el *beige* como categoría 12. En el estudio realizado participaron 20 sujetos (10 hombres y 10 mujeres) y se usaron 446 muestras. A los sujetos participantes se les aplicó el test de Ishihara [56] para detectar daltonismo. Luego de realizar el experimento, a las conclusiones de Boynton y Olson, ellos agregaron estas:

- No hay diferencias significativas entre hombres y mujeres en cuanto al número de términos básicos y no básicos usados.
- Mejor representación de los colores en el espacio Munsell con respecto al OSA, debido a las limitaciones en rango de este último.
- Los colores conocidos como "landmark" (el azul, verde, rojo y amarillo) tienden a ser nombrados más rápidamente que los otros básicos cromáticos (púrpura, anaranjado y café) aunque la diferencia en tiempo de respuesta es pequeña.
- Los resultados del experimento usando términos no básicos sugieren que los términos *beige* y *crema* podrían resolver los problemas del nombramiento de color en regiones entre blanco, amarillo, anaranjado, rosado y café.
- Aunque en el sistema Munsell se logró una mejor representación del color, en opinión de los autores, todos los sistemas de color existentes tienen limitaciones para representar la complejidad del proceso de percepción del color.

Los dos trabajos anteriores se dedicaron a la localización de las categorías básicas en un espacio de color, observándose muestras del espacio que no reciben ninguna categoría. La solución dada tiene un carácter discreto. En estos trabajos tampoco se sugiere la implementación del resultado para alguna aplicación. En 1978 se realizó un estudio muy interesante por McDaniel y Kay [28] en el que se intenta generar un modelo de nombramiento del color que establecía una relación entre mecanismos neurofuncionales durante la asignación de colores y los términos lingüísticos usados. El modelo se basaba en la teoría de conjuntos difusos. La propuesta indica que cada categoría de color está descrita por una función que determina el grado de membresía a ella. A continuación se explican las principales ideas de la propuesta:

- Debido a que los procesos neurofisiológicos sobre los que yacen las categorías de color están basados en funciones continuas (ver Figura 2.10), un formalismo no discreto de

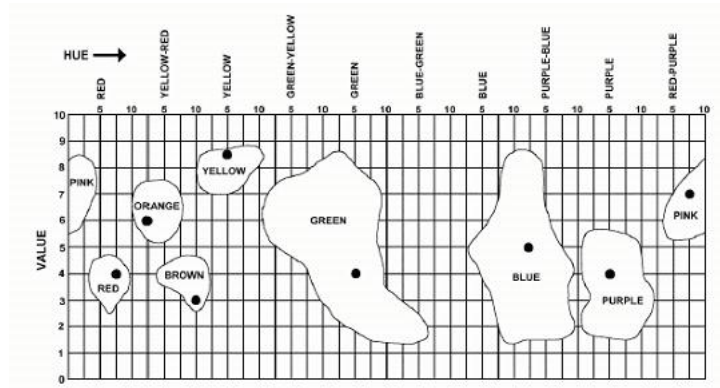


Figura 2.9. Zonas de consenso obtenidas por Sturges y Whitfield en el espacio de color de Munsell.

la categorización del color da una descripción adecuada del problema. Construcciones como azul claro, azul verdoso, rojo claro o rojo intenso son típicas de cualquier lenguaje y demuestran que los humanos reconocemos las categorías de color teniendo en cuenta grados de pertenencia.

- Las regiones, detectadas por Berlin y Kay, se pueden interpretar como aquellas donde las funciones de membresía de las categorías básicas alcanzaron su máximo valor.
- La representación difusa de las categorías básicas del color se forman a partir de 4 respuestas fundamentales del sistema visual humano (rojo, verde, amarillo y azul). Estas respuestas fueron obtenidas a través de datos de observadores humanos por Wooten en 1970 [57]. En la Figura 2.10 se muestran las respuestas espectrales mencionadas.

Los puntos B, R, Y y G tienen los mismos valores de longitud de onda que los de los focos encontrados en el experimento de Berlin y Kay. Obsérvese la declinación gradual en cada curva al aumentar la distancia respecto al foco en cualquier dirección. Aunque el modelo de McDaniel y Kay fue novedoso por su enfoque difuso, presentó desventajas: además de no darse el camino de su implementación, el modelo sólo considera 4 conjuntos difusos (rojo, verde, azul y amarillo) ; no está claro cómo definir el resto de las categorías básicas y como hacer que el modelo incorpore el conocimiento de otros colores [5].

Otro ejemplo de trabajo basado en la neurofisiología del color es el propuesto por Cairo [27], donde se elabora un modelo en 4 dimensiones: longitud de onda, intensidad, pureza y estado de adaptación de la retina. En el modelo las 11 categorías básicas se definen como una combinación de esas cuatro dimensiones, además se propusieron 4 categorías más: *azul cielo*, *turqueza*, *limón*

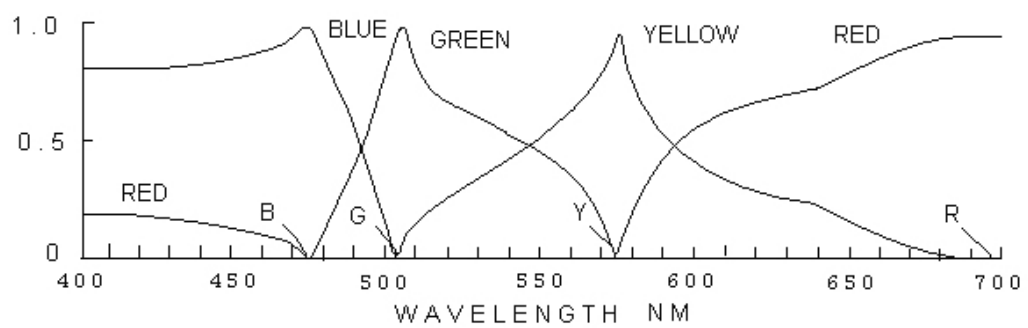


Figura 2.10. Respuestas espectrales, como funciones de membresía a las categorías roja, verde, azul y amarilla.

y *khaki*. Aunque este modelo resultó interesante por considerar el proceso de adaptación, el uso de parámetros físicos, en lugar de parámetros perceptuales hizo difícil su implementación al igual que en el modelo de McDaniel y Kay.

En 1994 Lammens [58] propuso el primer modelo paramétrico para el nombramiento del color usando diferentes espacios de color. En este modelo cada categoría está descrita por una función característica cuyos parámetros se calculan a partir de los datos obtenidos por Berlin y Kay en el espacio Munsell. A continuación se dan algunos detalles de la propuesta:

- Se crea un nuevo espacio de color, el cual denominan espacio neuropsicofísico (NPP) por usar los resultados de De Valois y sus colegas[4]. Los datos obtenidos por Berlin y Kay (focos y regiones de cada categoría) se llevan a este nuevo espacio.
- Como función para modelar cada categoría se elige una función Gaussiana. La selección de esta función está justificada por estudios realizados por Sherpard [59] en 1987 en psicología; él argumentó que la probabilidad de generalización de una categoría existente ante un nuevo estímulo, es una función monotónica de la distancia normalizada en el espacio psicológico del estímulo no conocido, con respecto al conocido que está en la categoría, posteriormente especificó que esta función puede ser aproximada por una exponencial y que esta podía ser una Gaussiana. Así el modelo propuesto para cada categoría viene dado por:

$$G_n(x) = e^{\frac{1}{2} \left(\frac{\sqrt[2]{\sum_{i=1}^N (x_i - \mu_i)^2}}{\sigma} \right)^2}. \quad (2.33)$$

En la expresión (2.33) μ se interpreta como la localización del centro o foco de la categoría, y el valor de la función como el grado de pertenencia de un estímulo x a una categoría. Con el valor de σ se controla la porción del volumen del espacio de color que se incluye en la categoría. Obsérvese que esta función tiene propiedades interesantes [58]: el máximo se encuentra en μ ; sus valores decrecen monotónicamente a medida que se alejan de μ ; una fácil definición matemática usando sólo dos parámetros μ y σ . La función es positiva excepto en distancias infinitas de μ [58].

Para el cálculo de los parámetros de cada categoría se minimiza el siguiente funcional:

$$E = \sum_{i=1}^N [G_n(\tilde{v}_i, \tilde{s}\tilde{\mu}, \sigma) - \theta]^2 + \varpi_f \sum_{i=1}^{N_f} [G_n(\tilde{f}_i, \tilde{s}\tilde{\mu}, \sigma) - 1]^2 + \sum_{i=1}^{N_0} [T(G_n(\tilde{o}_i, \tilde{s}\tilde{\mu}, \sigma))]^2. \quad (2.34)$$

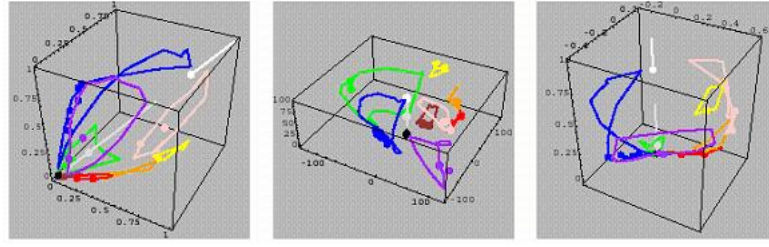


Figura 2.11. Los contornos (polígonos) y los focos (puntos) de las categorías básicas detectadas por Berlin y Kay, representadas en los espacios XYZ, L*a*b* y NPP (de izquierda a derecha).

En la expresión (2.34) cada término expresa lo siguiente:

E: es la suma total del error cuadrático para una categoría.

Primer término: indica el error respecto a los vértices del polígono (ver Figura 2.11) que delimita las regiones detectadas por Berlin y Kay siendo:

N_v : número de vértices del polígono.

\tilde{v}_i : vector que define vértices del polígono.

\tilde{s} : vector de escalamiento para μ .

σ : parámetro que define ancho de G_n .

θ : umbral que define la pertenencia a una categoría (los autores eligieron $\theta = 0.5$)

Segundo término: cuantifica el error respecto a los puntos focales, donde:

w_f : indica peso del error (los autores asumen $w_f = \frac{1}{3}$)

N_f : número de focos

\tilde{f}_i : vector que describe el foco.

Tercer término: cuantifica el error con respecto a otros representantes de la categoría, donde:

N_o : número de otros representantes.

T : función de umbral

$$T(x) = \begin{cases} 0, & x \leq \epsilon; \\ x, & \text{en otro caso.} \end{cases} \quad (2.35)$$

o_i : vector de los otros representantes de la categoría.

Los *otros representantes* se eligen tomando los puntos más cercanos del borde o del foco. La idea de introducir este término es mover las fronteras creadas por los datos iniciales (reajuste de las fronteras iniciales). El proceso de minimización se realiza respecto a σ y a \tilde{s} para cada componente de estos vectores. Con el primer término se trata de mover los bordes de la región que delimita una categoría según θ ; con el segundo término se tiene como objetivo lograr que

en los puntos focales el modelo de valores lo más cercanos posibles a 1. Una vez determinados los valores de σ y \tilde{s} , se procede a un segundo paso donde se distorsiona la forma obtenida de las fronteras en el paso anterior (paso 1), en el que por la métrica usada (Euclidiana) tiende a obtenerse contornos en forma de esfera. En el paso 2 se minimiza el funcional dado por la expresión (2.36).

$$E = \sum_{i=1}^{N_v} [G_n(\tilde{s}' \tilde{v}_i, \tilde{s}' \tilde{\mu}', \sigma') - \theta]^2 + \varpi_f \sum_{i=1}^{N_f} [G_n(\tilde{s}' \tilde{f}_i, \tilde{s}' \tilde{\mu}', \sigma') - 1]^2 + \sum_{i=1}^{N_0} [T(G_n(\tilde{\sigma}' \tilde{o}_i, \tilde{\sigma}' \tilde{\mu}', \sigma'))]^2. \quad (2.36)$$

La expresión (2.36) es similar a la expresión (2.34), solo que ahora tanto datos como focos son escalados por \tilde{s}' . El valor de $\tilde{\mu}' = \tilde{s} \tilde{\mu}$.

Una vez estimados los parámetros de cada modelo, la asignación a una categoría queda determinada por:

1. La evaluación en cada modelo el punto de color.
2. La elección de aquellas salidas mayores que cierto umbral.
3. El ordenamiento de los resultados obtenidos en 2.
4. La asignación del nombre asociado al mayor valor en 3.

El modelo propuesto por Lammens resultó ser el primero de tipo paramétrico. Éste fue implementado en el espacio RGB, en el XYZ y el L*a*b*, y los resultados fueron comparados con los obtenidos en el espacio creado en la propuesta: el NPP; llegándose a la conclusión que el mejor desempeño estaba en el espacio L*a*b*. La propuesta no considera modificadores de intensidad y saturación y a pesar de lo novedoso del trabajo en su tiempo no ha sido de uso extensivo en aplicaciones de visión.

Otro enfoque del problema del nombramiento del color y la descripción de la composición del color es el dado en [5]. El trabajo presenta un modelo computacional para la categorización del color usando la nomenclatura definida por "National Bureau of Standard" [60], donde el léxico está definido por:

- Matiz (hue): 28 nombres contruidos a partir del rojo, anaranjado, amarillo, verde, azul, violeta, púrpura, rosado, café, olivo, negro, blanco y gris.
- Brillo: muy oscuro, oscuro, medio, brillante y muy brillante.
- Saturación: grisoso, moderado, fuerte y vivo.

- Brillo/saturación: brillante, pálido.

Como aspectos interesantes de la propuesta está el hecho de considerar no sólo las categorías fundamentales de color sino también éstas con sus modificadores de brillo y de saturación, lo que hasta el momento otros autores no habían considerado. El otro punto interesante es la medida de similitud creada.

En la elaboración del modelo se desarrollaron 3 experimentos (listado de colores, descripción de la composición del color en imágenes, nombramiento del color). En los experimentos participaron 10 sujetos. Las conclusiones fueron las siguientes:

- Los sujetos listan 11 colores básicos coincidiendo con los fijados por Berlin y Kay, salvo algunas excepciones que incluyen el beige, el violeta y el cyan. No se nombran más de 14 colores.
- Durante la descripción del color se emplean los mismos términos que al listar los colores importantes y se usan los términos de saturación y brillo.
- En la mayoría de los casos se describen regiones bien definidas en color, siendo muy difícil la descripción de colores muy oscuros, que generalmente existen en la naturaleza debido a sombras o cambios de iluminación.

Como en otras investigaciones acerca del sistema visual humano, con estos experimentos el autor confirma que nuestro sistema visual realiza un promedio espacial que explica la forma en que se percibe el color; este grado de promediación depende de las frecuencias espaciales presentes y de la interacción entre colores, del tamaño de los objetos y el contexto global. Ejemplo de lo dicho se tiene en las texturas las cuales se describen con pocos colores; los bordes por su parte no se promedian, y creamos la percepción de un solo color en regiones uniformes. Es por todo lo dicho anteriormente que la percepción humana puede considerarse como un filtrado pasa bajo adaptativo, lo que se debe tener en cuenta en cualquier aplicación relacionada con la percepción del color si se quiere lograr resultados similares al juicio humano [5]. Usando las muestras para las cuales se obtuvo consenso durante el experimento de nombramiento de color, se crea una métrica usando los espacios $L^*a^*b^*$ y HSL. Con ella se clasifica cualquier otro nuevo estímulo de color (c_x) respecto a un prototipo (c_p). La métrica propuesta viene dada por la expresión:

$$D(c_p, c_x) = D_{Lab}(c_p, c_x)[1 + k(D_{Lab}(c_p, c_x))\Delta d(c_p, c_x)]. \quad (2.37)$$

D_{Lab} es la distancia euclidiana en el espacio $L^*a^*b^*$ y el valor de Δd está en función de la distancia en el espacio HLS. Ambas distancias dan una medida de la disimilaridad en los espacios correspondientes; el coeficiente $k(\cdot)$ se introduce para evitar modificar distancias entre puntos muy cercanos y limitar el incremento para distancias muy grandes ($k(l) = 0$, si $l < 7$ y $k(l) = \text{const}$ si $l > 30$). Los autores validaron su modelo comparando los resultados de su métrica con observaciones humanas y se encontró un 91 % de acuerdo. En el artículo se propone aplicar el modelo para nombramiento de color de las regiones extraídas por un método de segmentación. En el siguiente trabajo, elaborado por Robert Benavente y colaboradores [61] se desarrolla un modelo paramétrico, siguiendo la idea desarrollada por Lammens [58]. El modelo del nombramiento de color se basa en la teoría de conjuntos difusos. En calidad de función de membresía para indicar el grado de pertenencia a cada una de las categorías básicas se usan las funciones sigmoide y Gaussiana. Los parámetros de las funciones se calculan a partir de datos obtenidos por observaciones humanas. El modelo se evalúa en el espacio Munsell. A continuación los principales detalles del trabajo:

- En el experimento participan 10 sujetos, a los cuales se les muestran 422 muestras dos veces. Para asignar nombres, cada sujeto distribuye 10 puntos entre 11 categorías. En caso de seguridad en un solo nombre esa muestra toma la máxima puntuación. Para cada muestra se promedia la puntuación obtenida y se normaliza entre 0 y 1. De esta forma se tiene una asignación fuzzy hecha por humanos. Con dicha asignación se obtiene un descriptor de cada muestra de color (CD) observada.
- Para cada categoría se estiman los parámetros de su función de membresía mediante el siguiente funcional:

$$\min \frac{1}{2} \sum_{j=1}^s (f_{c_k}(x^j, \theta_k) - CD_k(x^j))^2. \quad (2.38)$$

donde:

S : número de muestras.

C_k : categoría de color.

X_j : muestra j .

$CD_k(X_j)$: la componente k del descriptor de color obtenido para la muestra X_j .

θ_k : parámetros de la función de membresía f_{C_k} .

- Se crea un espacio denominado uvI, cuyas dos primeras componentes representan color y la tercera intensidad. Sobre el espacio uvI se desarrolla el experimento y en aras de comparación se usa el espacio $L^*a^*b^*$.

Como modelo de función de membresía para las 11 categorías se tomó la función Gaussiana, al igual que Lammens en 1994; por esta razón los parámetros a estimar en la expresión (2.38) son $\tilde{\mu}$ (vector de medias estimado) y $\tilde{\Sigma}$ (la matriz de varianza covarianza estimada) de cada Gaussiana asociada a una categoría. Una vez estimados los parámetros se analizó el error cuadrático para cada modelo y se llegó a la conclusión de que la función Gaussiana era una buena aproximación para las categorías acromáticas, en cambio para el caso acromático no se había conseguido un buen ajuste, esto dio lugar a usar otro modelo para el caso cromático y la decisión fue usar una combinación de una función sigmoide con una gaussiana. En una sola dimensión la expresión es:

$$GS(x, \beta, \mu, \sigma) = \frac{1}{1 + e^{-\beta x}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}. \quad (2.39)$$

El tratamiento del caso cromático no termina con la expresión (2.39). Teniendo en cuenta que las categorías no aparecen en todos los niveles de intensidad, el espacio se divide en 3 niveles (20 % y 40 % del máximo de intensidad). Luego de esta división, el proceso de estimación de parámetros para el caso cromático se hace considerando las muestras correspondiente a cada nivel de intensidad. Terminada la modelación, la asignación de nombre de color se realiza acorde a la siguiente expresión:

$$N(\tilde{x}) = t_k | m_k = \max f_{ci}(\tilde{x}). \quad (2.40)$$

$N(\tilde{x})$ es la función que asigna uno de los 11 nombres de color (t_k) a la muestra \tilde{x} ; los autores consideran además que sólo cuando el valor máximo obtenido es mayor que 0.5 se realiza la asignación de nombre, en caso contrario la incertidumbre es muy grande.

El comportamiento de la propuesta se evalúa con respecto a los resultados obtenidos por Berlin y Kay viéndose un buen desempeño. Asimismo se observó que los resultados obtenidos en el espacio $L^*a^*b^*$, no fueron mejores que los obtenidos en el espacio creado (uvI). No se hace propuesta de aplicación.

Un trabajo interesante es el reportado en [62] por M. Seaborn y colaboradores, en el cual se crea una medida de similitud entre colores y se propone un espacio de color basado en un enfoque difuso. El número de clases que consideran en su método es también 11 usándose los términos básicos de Berlin y Kay; y los datos obtenidos por Sturges y sus colaboradores en [29]. Como detalles de la investigación está la distinción que hacen de la función de membresía respecto a categorías cromáticas y de la de membresía para las acromáticas. Las funciones de membresía usan la información de las componentes H, V y C del espacio Munsell. En [62] se señala que H y V determinan la pertenencia a las categorías cromáticas, y que la membresía a categorías

acromáticas está dada sólo por la información en V. Además de lo dicho anteriormente señalaron que la relación entre cromáticas y acromáticas está determinada por el cromatismo. A partir de aquí la expresión para la función de membresía de las clases cromáticas es:

$$\mu_{\mathcal{R}_r}(\tilde{i}_l^{HV}) = \frac{1}{\sum_{q=1}^R (\|\tilde{i}_l^{HV} - \tilde{v}_r^{HV}\| / \|\tilde{i}_l^{HV} - \tilde{v}_q^{HV}\|)^{2/(m-1)}}; 1 \leq r \leq 8. \quad (2.41)$$

donde r es el número de categorías cromáticas, \tilde{i}_l^{HV} está formado por las coordenadas H y V de un punto en el espacio Munsell, \tilde{v}_r^{HV} es el prototipo de la categoría r . La función de membresía para el caso acromático es:

$$\mu_{A_a}(\tilde{i}_l^V) = \frac{1}{\sum_{q=1}^A (\|\tilde{i}_l^V - \tilde{v}_a^V\| / \|\tilde{i}_l^V - \tilde{v}_q^V\|)^{2/(m-1)}}; 1 \leq a \leq 3. \quad (2.42)$$

En [62] se propone además una medida de similitud basada en los trabajos desarrollados por Tversky en [63]. Esta medida la comparan con la distancia euclidiana, llegando a la conclusión de que con la medida propuesta se logra un mayor acuerdo con el juicio humano al analizar similitud entre colores. Se realizan comparaciones con los espacios RGB, L*a*b*, L*u*v* y el HSV, destacándose el espacio HSV como muy bueno para detectar similitud para el caso de pequeñas y grandes distancias. Con el enfoque difuso realizado en este trabajo se logra representar un color como una colección de funciones de membresía a las 11 categorías básicas definidas por Berlin y Kay.

Entre las aplicaciones de los trabajos estudiados se encuentran el uso en CBIR (*content based image retrieval*). El sistema Pisaro [64] es un ejemplo; aquí se introduce la idea del histograma de categorías perceptuales. Dicho histograma consta de 11 bins, cada uno de los cuales se corresponde con una de las categorías básicas. Cada imagen en la base de datos se representa usando este histograma y las respuestas a las consultas, que usan como característica el color, se hacen atendiendo a la proporción de colores dada por el histograma en cada imagen. En [65] se propone el sistema CHROMA, en el cual se usan 10 colores basados también en los trabajos de Berlin y Kay. La propuesta consiste en que la proporción de estos colores en las imágenes es usada para crear una estructura jerárquica basada en la dominancia de cada una de las categorías en las imágenes.

En [5] se da una aplicación del nombramiento de colores como paso final de la segmentación. Luego de realizar un preprocesamiento a la imagen, se realiza la segmentación usando *mean shift* [22] y las regiones de color encontradas se nombran según modelo en [5]. En [58] se propone el uso de los modelos de categorización para la robótica, usando el color como guía en el trabajo

del robot.

2.4. La segmentación en colores y los modelos de categorización del color

En la sección (2.2) se comentó sobre algunas de las vías usadas para la segmentación: umbralado por histogramas, agrupamiento en un espacio de características, métodos basados en regiones, métodos basados en bordes, redes neuronales, métodos difusos y la combinación de técnicas Bayesianas con campos aleatorios Markovianos. Cada método comentado tiene ventajas y desventajas siendo muy promisorios el enfoque difuso y la combinación de técnicas Bayesianas con campos aleatorios Markovianos pues éstas modelan de una forma natural la incertidumbre presente en imágenes reales y por tanto reflejan el carácter perceptual de la segmentación. En todos los métodos comentados hay dos problemas comunes:

- ¿Cómo encontrar cuales y cuantos colores son los más significativos en la imagen?
- ¿Cómo procesar la información de color?

El primer aspecto está relacionado con la solución a la tarea de validación de grupos y determina la cantidad de regiones resultantes en una imagen pudiendo ocurrir dos casos:

- sobresegmentación: el número de grupos o clases obtenidos es mayor que el necesario.
- omisión de clases.

El segundo problema común involucra la selección de una representación del color y como considerar sus componentes.

En la sección (2.3), por su parte, se discute sobre diferentes modelos de categorización, que explican el proceso de nombramiento del color por observadores humanos y tienen como punto de partida las investigaciones hechas por Berlin y Kay [6] sobre el lenguaje usado al asignar colores, en diferentes idiomas. Acorde a las investigaciones hechas en los trabajos revisados, existe un número de 11 categorías básicas del color, estas son: rojo, verde, azul, amarillo, anaranjado, café, rosa, púrpura, gris, negro y blanco. A partir de ellas se realiza la partición de un espacio de color. En la sección (2.3) se pudo apreciar también, como es procesada la información de color: todas las componentes del espacio seleccionado, se consideraron al mismo tiempo, independientemente de la representación del color usada.

Entre los modelos revisados [58], [61], [62] se propone describir la posición de un voxel en el espacio, no sólo por sus coordenadas, sino también por un vector, cuyas componentes indican el grado de pertenencia a cada categoría básica. Se tiene entonces que los modelos de categorización permiten particionar un espacio de forma no arbitraria y describen, de forma similar al observador humano, el proceso del nombramiento de color. Lo dicho anteriormente indica que en la fusión de los modelos de categorización con la segmentación en colores podría encontrarse una vía para resolver los problemas antes planteados, por otro lado si la fusión se realiza con métodos basados en técnicas Bayesianas de interacción espacial se lograría simular en parte el comportamiento del sistema visual humano y por tanto el carácter perceptual de la segmentación estaría incluido en el método creado.

En esta tesis se propone un nuevo modelo de categorización del color y se fusiona el mismo con técnicas Bayesianas y campos aleatorios Markovianos para dar lugar a una segmentación lingüística del color.

Capítulo 3

Sobre la propuesta de tesis

Los modelos de categorización del color se basan en los estudios de la percepción humana (sección (1) y sección (2)). A partir de la investigación realizada por Berlin y Kay en 1969 [6] fueron propuestos varios modelos que simulan el proceso de nombramiento del color. Los primeros trabajos dieron como resultado una partición dura del espacio de color seleccionado [6],[26],[29]. Mas adelante se propusieron modelos en los que se consideró la incertidumbre presente durante la asignación de colores [28],[62],[61], [5]. En la generalidad de los trabajos reportados se usan las 11 categorías básicas definidas por Berlin y Kay por ser estas salientes perceptuales de información; aunque se reportan investigaciones con otro vocabulario [5],[60]. La definición de prototipos de los colores usados en el vocabulario seleccionado sigue siendo aún un problema crucial, mas aún cuando el modelo creado se basa en una medida de similitud con respecto a los prototipos. Las particiones de los espacios, reportadas hasta el momento, no describen la geometría de la región ocupada por cada categoría y no siempre es posible conocer la categoría de color que le corresponde a cada voxel [61]. La elaboración de un modelo, que permita incorporar nuevas clases de color y de esta manera hacerlo adecuado para su aplicación sigue siendo también una tarea a resolver.

Basado en estos aspectos en esta tesis se elabora un nuevo modelo de categorización del color, en el que cada voxel del espacio de color queda descrito por un vector de probabilidades, cuyas componentes expresan el grado de pertenencia a cada categoría. A través del vector obtenido se calcula la partición del espacio y se conoce la ubicación y geometría aproximada de cada clase de color. Los vectores de probabilidades obtenidos dan lugar además, a un diccionario probabilístico. A diferencia de otros modelos reportados no es necesario conocer los prototipos de cada categoría y la estructura del mismo permite ampliar el vocabulario si fuese necesario. En la sección (2) se expresó lo interesante que sería vincular los modelos de categorización

con la segmentación de imágenes en colores para resolver los problemas del número de clases a considerar y el procesamiento de la información de color como un todo. Aunque en [5] se reporta el uso del modelo creado para nombramiento de regiones de color después de la segmentación, aun no se registra la aplicación directa de un modelo de categorización del color para llevar a cabo la segmentación.

En este trabajo de tesis se propone usar el modelo elaborado para la segmentación. Considerando el diccionario probabilístico obtenido y un enfoque Bayesiano combinado con técnicas de interacción espacial se construye un nuevo método de segmentación proponiéndose una generalización del trabajo realizado en [30]. Con la nueva propuesta, se simula la incertidumbre presente en zonas de color muy oscuras, así como la interacción entre colores que el observador humano realiza al describir una escena en colores. Además de describir cada región de color extraída por una de las 11 categorías básicas, se consideran los atributos de intensidad *claro* y *oscuro* para cada región encontrada. Finalmente el proceso de segmentación permite conocer en lenguaje natural, la composición en color de la imagen estudiada usando las categorías básicas de color.

A continuación, se explican las herramientas usadas y los aspectos teóricos de la propuesta.

3.1. El enfoque Bayesiano junto a técnicas de interacción espacial

La segmentación es un problema de agrupamiento, de ahí que algoritmos como k-medias, c-medias difusas e ISODATA [66] hayan sido exitosamente usados, sin embargo estos métodos sólo incluyen los datos de cada pixel y omiten la interacción espacial, con la cual se tiene información adicional muy útil para resolver la segmentación como problema mal planteado [30]. Una vía eficiente para segmentar, teniendo en cuenta las observaciones en cada pixel y su información de contexto, es la estimación Bayesiana junto a modelos de campo aleatorio Markoviano (CAM) [52],[53] [55], [54],[30]. Con este enfoque se calcula un campo de etiquetas (f), asumiéndose que la dependencia de probabilidades que determinan la etiqueta en un pixel, tiene lugar solo con respecto a los pixeles de su vecindad. Las restricciones a imponer sobre la solución se logran a través de la distribución a priori, la cual se considera Markoviana, y por tanto toma la forma de una distribución de Gibbs [54]:

$$P_f(f) = \frac{1}{Z_f} \exp \left[-\beta \sum_c V_c(f) \right]. \quad (3.1)$$

En la expresión (3.1) la suma se realiza sobre todos los cliques [53] del sistema de vecindad definido en la retícula L . En la nomenclatura que se usará a continuación r denota pixeles o sitios con coordenadas (x, y) y $< r, s >$ indica pares de pixeles vecinos. En el caso tridimensional r denota los voxels en el espacio con coordenadas (x, y, z) .

La función de potencial V_c se evalúa solamente sobre los cliques; β es un parámetro positivo que regula la influencia o fuerza de la función de potencial y Z_f es una constante de normalización. La selección de la función V_c depende del tipo de valores (reales o discretos) del espacio de estado de las etiquetas y de las restricciones que se le impone a f . Como parte de la estimación Bayesiana, es necesario modelar las observaciones ($g(r)$). Una forma posible es considerar el modelo dado en la expresión (3.2) [55],[67].

$$g(r) = \sum_k \Phi(r, \theta_k) p_k(r) + \eta(r). \quad (3.2)$$

Bajo este modelo en (3.2) se asume que la imagen está compuesta por k regiones (R_k). En (3.2) $\eta(r)$ denota el ruido que distorsiona las observaciones. El mismo tiene distribución conocida P_η , asumiéndose que para todo $\eta(r), r \in L$ las variables $\eta(r)$ son independientes, tienen media 0 y desviación standard σ , comportándose como una distribución Gaussiana. Por su parte $\Phi(.,.)$ es un modelo paramétrico que describe el campo f a estimar y define las regiones R_k en la imagen. Ejemplos de funciones usadas para modelos paramétricos son los splines (ver apéndice A) y los polinomios; θ_k es el vector de parámetros asociado a cada R_k en la imagen; $p_k(r)$ indica el grado de pertenencia del pixel r a la clase k y satisface la siguiente expresión:

$$\sum_k p_k(r) = 1, \forall r \in L \quad p_k(r) \geq 0, \forall k, r. \quad (3.3)$$

Si p fuera binaria, se estaría en el caso discreto y entonces la estimación del campo f se haría más compleja. Un caso particular de (3.2) es cuando el modelo que describe cada región es constante para cada sitio. La ecuación (3.4) describe este caso.

$$g(r) = \sum_k \mu_k p_k(r) + \eta(r). \quad (3.4)$$

En (3.4) el modelo que describe a R_k es μ_k .

Considerando el ruido ($\eta(r)$) como una variable aleatoria independiente (ruido blanco) se puede construir un campo de distribuciones condicionales independientes: $P_{g_r|f_r} = D_r, r \in L$. Este campo se denomina campo de verosimilitud. Finalmente, la distribución a posteriori es calculada según regla de Bayes:

$$P_{f|g} = \frac{P_{g|f}(f)P_f(f)}{P_g(g)} = \frac{1}{Z} \exp^{-U(f)}. \quad (3.5)$$

En (3.5) $U(f)$ se define mediante la expresión:

$$U(f) = - \sum_r D'_r(f_r) + \beta \sum_{\langle r,s \rangle} V_{rs}(f_r, f_s). \quad (3.6)$$

$D'_r(f_r)$ viene dado por el logaritmo de la función de verosimilitud (D_r) y mide la correspondencia entre los datos (g) y el campo estimado (\hat{f}). Para calcular el estimador Bayesiano óptimo (\hat{f}) es necesario definir una función de costo que depende de la \hat{f} estimada y de los valores de f . El estimador óptimo es calculado entonces minimizando el valor esperado de la función de costo definida, respecto a la distribución a posteriori [68], [54]. Si se considera la función de costo dada en (3.7) y se calcula el valor esperado de la misma (3.8) se tiene que, minimizar el valor esperado de la función de costo definida en (3.7) es equivalente a maximizar la probabilidad a posteriori (3.10). Lo dicho anteriormente implica minimizar la función de energía U . Este estimador se conoce como MAP (Maximum a Posteriori)[54].

$$C(f, \hat{f}) = 1 - \delta(f - \hat{f}), \quad (3.7)$$

$$E[C(f, \hat{f})] = \int_{\Upsilon} C(f, \hat{f}) P_{f|g} df, \quad (3.8)$$

$$= \int_{\Upsilon} (1 - \delta(f - \hat{f})) P_{f|g} df, \quad (3.9)$$

$$= 1 - P_{\hat{f}/g}. \quad (3.10)$$

En (3.8) Υ representa el espacio de estado de todas las configuraciones posibles del campo f . La función de costo definida en (3.7) puede no ser apropiada para el caso discreto, pues solo asigna costo 0 a la solución ideal y 1 para cualquier otro caso, sin tener en cuenta el número de píxeles clasificados erróneamente. Esto explica el porque este estimador no es apropiado en imágenes muy ruidosas. Una solución para esto es la función de costo definida en (3.11), en la que se considera el número de píxeles donde hubo error [54].

$$C(f, \hat{f}) = \sum_{r \in L} [1 - \delta(f_r - \hat{f}_r)]. \quad (3.11)$$

En este caso el valor esperado de la función de costo se define como en (3.13):

$$\sum_{f \in \Upsilon} C(f, \hat{f}) P_{f|g} = \sum_{f \in \Upsilon} \sum_{r \in L} [1 - \delta(f_r - \hat{f}_r)] P_{f|g}(f), \quad (3.12)$$

$$= |L| - \sum_{r \in L} \sum_{f: f_r = \hat{f}_r} P_{f|g}(f) = |L| - \sum_{r \in L} \pi_r(\hat{f}_r). \quad (3.13)$$

En (3.13) se tiene que:

$$\pi_r(q) = \sum_{f: f_r = q} P_{f|g}(f). \quad (3.14)$$

En la expresión (3.14) se indica la distribución marginal a posteriori del pixel r . El estimador óptimo, dado por la función de costo en (3.11), se conoce Maximizador de las Marginales a Posteriori(MPM, Maximizer of the Posterior Marginals). La fórmula obtenida en (3.13) permite una solución exacta, sin embargo computacionalmente es muy costosa por el número de términos que ella involucra. Una vía para encontrar las marginales a posteriori es la que permiten los métodos estocásticos como el muestreador de Gibbs o el de Metropolis [52]. Bajo este enfoque las probabilidades marginales a posteriori se calculan contando el número de veces que cada elemento $\xi_r, r \in L$ del espacio de estado, toma un valor en particular q_k :

$$p_r(k) = \frac{1}{T_0} \sum_{t=1}^{T_0} \delta(\xi_r^t - q_k) \approx \sum_{f: f_r = q_k} P_{f|g}(f). \quad (3.15)$$

El problema de estos métodos consiste en determinar el número de veces óptimo (T_0), que es necesario tener, para lograr una buena aproximación de las marginales a posteriori. En el trabajo desarrollado por M. Rivera, J. Marroquin y O.Oceguera en [30] se da un método para el cálculo aproximado de las distribuciones marginales a posteriori, donde se producen resultados similares a los dados por la vía estocástica y en menor tiempo. En esta tesis se hará uso del trabajo en [30] y se propone una generalización del mismo.

3.2. Campos aleatorios de medida Markovianos Gaussianos con entropía controlada

El modelo GMMF (*Gauss-Markov Measure Field*) [54] estima distribuciones marginales con una alta entropía [30], si las mismas se comparan con las halladas asintóticamente por Monte Carlo Markov Chain (MCMC); estas últimas son consideradas como verdaderas. El objetivo del modelo GMMF con entropía controlada es obtener un modelo GMMF que de cómo resultado

marginales empíricas con igual comportamiento al de las verdaderas a través de un cálculo eficiente de las mismas [30]. En el caso del modelo GMMF se tiene el siguiente funcional:

$$U(p) = \sum_{r \in L} |p_r - \hat{p}_r|^2 + \lambda \sum_{\langle r, s \rangle} |p_r - p_s|^2. \quad (3.16)$$

En (3.16) \hat{p}_r es la verosimilitud normalizada. Con este modelo se tiene que las distribuciones marginales empíricas deberán ser similares a la verosimilitud de las observaciones y variarán suavemente en la retícula. El grado de suavidad depende del parámetro λ . Una vez estimado el campo p el estimador óptimo \hat{f}_r se calcula como la moda de cada vector de probabilidades p_r [54]. En GMMF se asume que en ausencia de información a priori el maximizador de (3.16) debe ser la verosimilitud normalizada. Dado que en GMMF la verosimilitud está determinada por un modelo Gaussiano la dependencia no lineal en el término de datos es alta, aspecto que dificulta su uso en procedimientos de tipo *Expectation-Maximization* (EM). Por otra parte la distribución obtenida para cada pixel (p_r) refleja alta entropía. Con el objetivo de eliminar estos problemas en [30] se introduce una nueva restricción, acorde con la cual se considera que la moda del estimador óptimo de las distribuciones marginales a posteriori coincide con el máximo de la verosimilitud (v). Esta restricción permite usar $\log v$, en lugar de v en el término de datos en la función U y de esta manera se tiene una dependencia cuadrática sobre los parámetros del modelo. El control de entropía se logra adicionando un término para penalizar distribuciones de alta entropía; para ello se usa el índice de Gini [68]: $-\sum_k \sum_r p_{kr}^2$. La función de energía es finalmente [30]:

$$U_{EC}(p, \theta) = \frac{1}{2} \sum_k \sum_r \left[p_{kr}^2 (-\log v_{kr}(\theta) - \mu) + \lambda \sum_{s \in N_r} (p_{kr} - p_{ks})^2 \right]. \quad (3.17)$$

sujeto a :

$$\sum_k p_{kr} = 1 \quad p_{kr} \geq 0, \forall k, r. \quad (3.18)$$

En (3.17) μ es un parámetro positivo que controla la entropía y $N_r = \{s : |r - s| = 1\}$ es el conjunto de vecinos más cercanos al pixel r . Para lograr que U sea cuadrática definida positiva debe cumplirse que μ sea pequeño. En la expresión (3.17) θ representa los parámetros de los modelos a usar.

Con el objetivo de incorporar las restricciones de igualdad dadas en (3.18) se construye el Lagrangiano:

$$L(p, \theta) = U_{EC}(p, \theta) - \sum_r \gamma_r (1 - \sum_k p_{kr}). \quad (3.19)$$

La minimización de (3.19) se efectúa usando un procedimiento de 2 pasos. Los estimadores MAP para las marginales (p) y para los parámetros de los modelos (θ) son calculados alternando la minimización de (3.19) respecto a p y θ respectivamente. Derivando (3.19) respecto a p , luego igualando a cero y despejando para p_{kr} y finalmente usando la restricción dada en (3.18) se obtiene la expresión siguiente para obtener la solución usando Gauss-Seidel [30]:

$$p_{kr} = \frac{n_{kr}}{m_{kr}} + \frac{1 - \sum_{l=1}^N \frac{n_{lr}}{m_{lr}}}{\sum_{l=1}^N \frac{m_{kr}}{m_{lr}}}. \quad (3.20)$$

En (3.20) $n_{kr} = \lambda \sum_{s \in |N_r|} p_{ks}$ y $m_{kr} = (-\log v_{kr} - \mu) + \lambda |N_r|$, siendo $|N_r|$, la cardinalidad del conjunto de vecinos del pixel r .

Una vez calculado p los modelos θ se determinan según (3.21):

$$\theta_{MAP} = \underset{\theta}{\operatorname{argmin}} \sum_k \sum_r \left[-p_{kr}^2 \log v_{kr}(\theta) \right]. \quad (3.21)$$

3.3. Generalización de Campos aleatorios de medida Markovianos Gaussianos con entropía controlada

En esta tesis se propone una generalización del modelo para campos de medida Markovianos con entropía controlada. A continuación se explican detalles.

Sea B , una matriz cuyos elementos indican la interacción o relación entre las componentes del vector p en (3.17):

$$B = [b_{kl}], \quad k = 1, \dots, n; l = 1, \dots, n. \quad (3.22)$$

Desarrollando el término de regularización en (3.17) se tiene que:

$$\frac{\lambda}{2} \sum_{s \in |N_r|} \sum_k (p_{kr} - p_{ks})^2 = \frac{\lambda}{2} \sum_{s \in |N_r|} \sum_k p_{kr}^2 + p_{ks}^2 - 2p_{kr}p_{ks} \quad (3.23)$$

Si se introduce la matriz dada en (3.22) y la expresión obtenida en (3.23) en (3.17) se obtiene:

$$U_{ECG}(p, \theta) = \sum_k \sum_r \left[p_{kr}^2 (-\log v_{kr}(\theta) - \mu) \right] - \lambda \sum_{\langle r, s \rangle} \sum_k \sum_l p_{kr} p_{ls} b_{kl} + \frac{\lambda}{2} \sum_k \sum_r |N_r| p_{kr}^2. \quad (3.24)$$

La expresión dada en (3.24) constituye una generalización de la propuesta elaborada en [30]. Esto conduce al siguiente problema de minimización:

$$\min_{(p,\theta)} U_{ECG}(p, \theta). \quad (3.25)$$

sujeto a :

$$\sum_k p_{kr} = 1 \quad p_{kr} \geq 0, \forall k, r. \quad (3.26)$$

Al construir el Lagrangiano se tiene:

$$L(p, \theta) = U_{ECG}(p, \theta) - \sum_r \gamma_r (1 - \sum_k p_{kr}). \quad (3.27)$$

La minimización de (3.27) es efectuada usando un procedimiento de dos pasos, al igual que en la minimización de (3.19). Derivando (3.27) respecto a p_{kr} , luego igualando a cero, despejando y finalmente usando la restricción dada en (3.18) se obtiene la expresión siguiente para obtener la solución usando un esquema Gauss-Seidel:

$$p_{kr} = \frac{\gamma_r - \lambda \sum_{s \in N_r} \sum_l p_{ls} b_{kl}}{2 \log v_{kr} + 2\mu - \lambda |N_r|}. \quad (3.28)$$

donde:

$$\begin{aligned} \gamma_r &= \frac{1 + \lambda S_1}{S_2}, \\ S_1 &= \sum_k \frac{1}{2 \log v_{kr} + 2\mu - \lambda |N_r|}, \\ S_2 &= \sum_k \frac{\sum_{s \in N_r} \sum_l p_{ls} b_{kl}}{2 \log v_{kr} + 2\mu - \lambda |N_r|}. \end{aligned}$$

El estimador MAP para los modelos θ se calcula según (3.21).

En (3.24) obsérvese, que cuando la matriz B es la identidad se tiene:

$$b_{kl} = \begin{cases} 1, & \text{si } k = l \\ 0, & \text{en otro caso.} \end{cases} \quad (3.29)$$

Para este caso se obtiene (3.17) como un caso particular.

En la sección (2.3) se detalló sobre el estado del arte de los modelos de categorización del color, haciéndose referencia al trabajo desarrollado por Boynton y Olson en [26]. Una de las conclusiones de los experimentos realizados en este trabajo fue que entre las categorías básicas del color existe una relación o enlace que se refleja en el hecho de nombrar un color usando mas de una de ellas. La no existencia de enlaces entre categorías es justificada por su

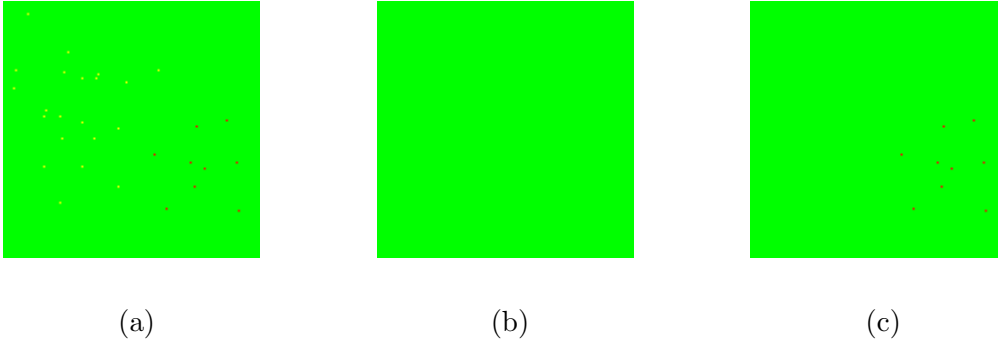


Figura 3.1. (a) Imagen original, (b) segmentación usando ecuación (3.18), (c) segmentación usando ecuación (3.24)

incompatibilidad perceptual. La relación entre categorías se expresó a través de una matriz de confusión, cuya ilustración gráfica aparece en la Figura 2.8 en la sección (2.3). Usando estos resultados experimentales, en este trabajo de tesis, se construye la matriz de interacción entre categorías de color (B_{11}), cuyos elementos se definen como se indica a continuación:

$$b_{kl} = \begin{cases} 1, & k = l; \\ -1, & k \# l \\ 0, & \text{en otro caso.} \end{cases} \quad (3.30)$$

En (3.30) $k \# l$ indica que entre las componentes k y l existe interacción, o sea entre la categoría k y la categoría l existe una conexión desde el punto de vista perceptual (ver Figura 2.8); $k, l \in \{1, 2, \dots, 11\}$, ya que se consideran las interacciones entre las 11 categorías básicas del color definidas por Berlin y Kay [6]. En la Figura 3.1 se muestra un ejemplo sintético, donde se compara el comportamiento de (3.17) y de (3.24). Usando los mismos parámetros ($\lambda = 1.5$, $\mu = 0.5$) la segmentación es diferente. Los colores verde y amarillo están relacionados perceptualmente (ver Figura 2.8), esto hace que en una región donde pixeles verdes y amarillos sean vecinos sus componentes correspondientes estén pesadas según (3.30) y la acción del término de regularización en (3.24) conlleva a eliminar la granularidad dada por los puntos amarillos, de forma similar a lo que hace nuestro sistema visual al promediar el color en una región. En cambio entre verde y rojo no hay relación (vea Figura 2.8), y los puntos rojos se mantienen. Si no se considera la interacción entre colores (ver Figura 3.24 b) la acción del término de regularización conlleva a eliminar cualquier región de color muy pequeña.

3.3.1. Estimación de la verosimilitud

Una componente importante de la estimación Bayesiana es la verosimilitud, que está presente en el término de datos del potencial propuesto en este trabajo (expresión (3.24)).

Con el objetivo de estimar la verosimilitud se cuenta con datos, obtenidos de un experimento de nombramiento del color. En el experimento se tienen N muestras de color y M sujetos. Cada sujeto, al observar una muestra de color (c), indica la categoría básica (k) más verosímil. Con este proceder la verosimilitud ($v = P(k|c)$) se determina experimentalmente como en (3.31):

$$v = P(k|c) \approx \frac{A_{kc}}{M}, \quad k \in \{1, 2, \dots, n\}. \quad (3.31)$$

En (3.31) A_{kc} indica el número de sujetos que asignaron la clase k a la muestra de color c y n denota el número de categorías de color usadas. Considerando la investigación hecha en [6], en este trabajo se usa $n = 11$.

Con el experimento descrito, solo es posible calcular la verosimilitud para un número de muestras limitado (N) de todo el espacio de color (Ω). Se necesita conocer la misma para todo Ω . Para lograr lo dicho anteriormente se proponen los siguientes caminos:

1. Uso de la estimación Bayesiana a través del modelo GMMRF (3.16) restringiendo el término de datos a $r \in D$, siendo D un subconjunto de voxeles (r), correspondientes a las muestras de color c ($D \subset \Omega$), de los cuales se conoce su verosimilitud experimentalmente ($v = P(k|c)$).
2. Uso de un modelo de splines cuadráticos, con un modelo de CAM para los coeficientes de los splines.

Los caminos propuestos permiten realizar una interpolación de la verosimilitud (v) y conocer la misma en todo Ω a partir de los datos del experimento de nombramiento de color.

Como resultado de esta estimación se obtiene un nuevo modelo de categorías del color en el que cada voxel del espacio, es descrito por un vector, cuyas componentes expresan el grado de pertenencia a cada una de las 11 categorías básicas de color. Usando este vector se conoce la categoría básica de color más probable para cada voxel, y por tanto se calcula la partición R en Ω como se expresa en (3.32)

$$R_r = \arg_{k=\{1,2,\dots,n\}} \max_k v_{kr}. \quad (3.32)$$

Finalmente se conoce la ubicación de cada una de las 11 categorías de color en el espacio Ω . El procedimiento seleccionado para el cálculo del campo de verosimilitud debe dar lugar a regiones conexas y suaves de cada categoría de color, de forma tal que la partición obtenida se corresponda con la percepción humana. Finalmente la solución encontrada al estimar el campo de verosimilitudes, es interpretada como un diccionario probabilístico, cuyo uso se verá más adelante.

Estimación de la verosimilitud a través de GMMRF

Al usar el modelo dado en (3.16) y restringiendo el término de datos a aquellos voxels (r) con información obtenida de los experimentos psicofísicos se estima la verosimilitud. La expresión dada en (3.16) bajo la condición dicha se convierte en (3.33):

$$E(\hat{p}(r)) = \sum_r |\hat{p}(r) - p(r)|^2 I(r) + \lambda \sum_{\langle r, s \rangle} |\hat{p}(r) - \hat{p}(s)|^2. \quad (3.33)$$

En (3.33) $I(r)$ es una función indicadora que se define como:

$$I(r) = \begin{cases} 1, & r \in D \\ 0, & \text{en otro caso.} \end{cases} \quad (3.34)$$

Las componentes del vector $p(r)$, para los voxels con información, están definidas como en (3.31), o sea este vector contiene la verosimilitud experimental.

Al derivar (3.33) con respecto a p_{kr} , igualando a 0 y despejando se tiene (3.35):

$$\hat{p}_{kr} = \frac{p_{kr} I(r) + \lambda \sum_{s \in |N_r|} p_k(s)}{(I(r) + \lambda |N_r|)}. \quad (3.35)$$

En la Figura 3.2 se ilustra la región asignada a la categorías *café y verde* en el espacio RGB luego del proceso de interpolación. Obsérvese que no se tienen regiones conexas para un mismo color, lo que influye negativamente en la composición del diccionario probabilístico a obtener y por tanto la modelación que se tiene no es la deseada. El comportamiento de la partición respecto a otras clases fue similar. La cantidad de datos disponibles (sólo el 0.002% de los voxels del espacio tienen conocida la verosimilitud a través del experimento de nombramiento de color) y la distribución de los mismos no fue suficiente para lograr la partición adecuada por esta vía. El tiempo de interpolación fue de 17h aproximadamente, usando los parámetros que se muestran en los ejemplos y un procesador Pentium IV de 1GHz y 1 GB RAM.

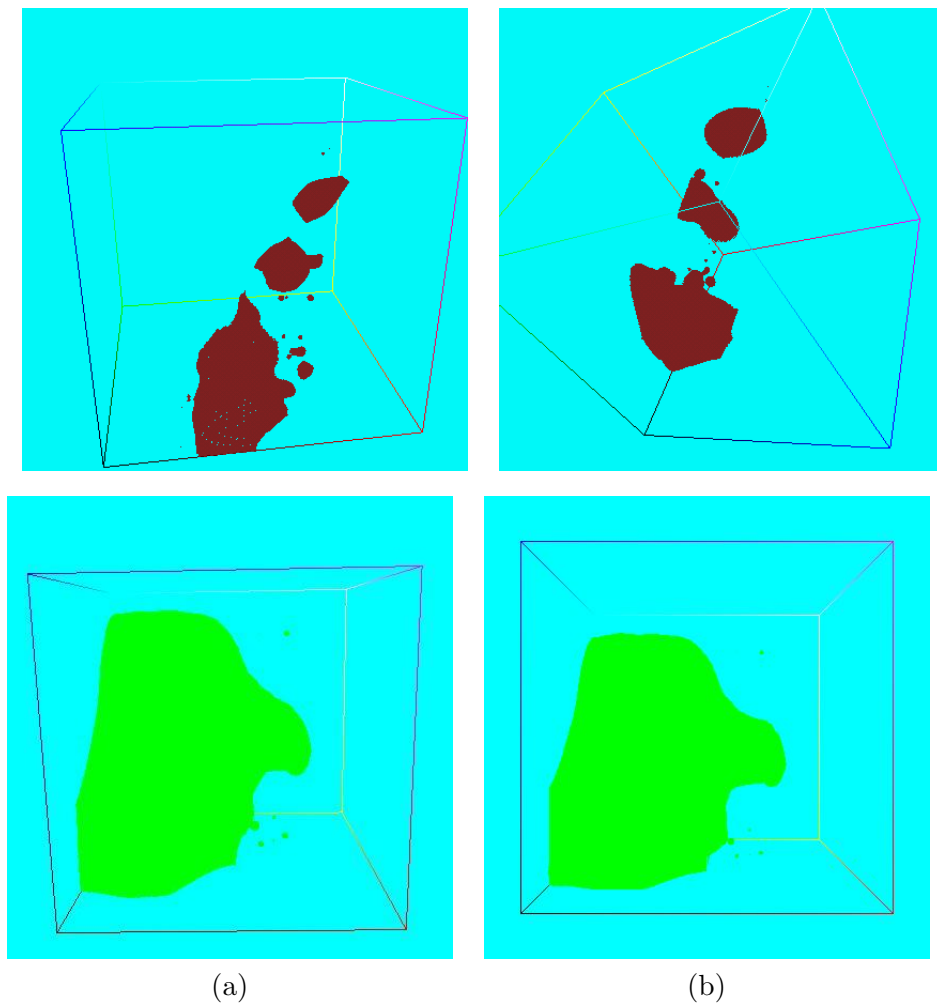


Figura 3.2. Interpolación por GMMRF, a) $\lambda = 0.05$ b) $\lambda = 10$. En ambos casos el número de iteraciones es 2000.

La necesidad de lograr una partición suave y regiones de color conexas, cuando estas estuviesen asociadas a una misma categoría, condujo a explorar la vía de interpolación basada en splines cuadráticos, ya que estos permiten lograr una buena aproximación de los datos, y por tanto una buena interpolación, incluso cuando se cuenta con poca información, como es el caso que se discute.

Estimación de la verosimilitud a través de splines cuadráticos

Teniendo en cuenta el buen desempeño de los splines cuadráticos (ver apéndice A) en tareas de interpolación, se explora este camino como otra vía para la estimación de la verosimilitud en todo Ω .

Cada voxel $r \in D$, tiene asociado un vector v como sigue:

$$v_r = \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{pmatrix} \quad (3.36)$$

En (3.36) cada componente del vector se calcula como en (3.31).

En el proceso de interpolación cada categoría se modela como una combinación lineal de splines cuadráticos:

$$v_k(r) = \sum_j \alpha_{kj} \beta_j(r). \quad (3.37)$$

En la expresión (3.37) $k \in \{1, 2, \dots, n\}$ indica la categoría que se modela. Los splines cuadráticos, que son funciones de soporte compacto, son colocados sobre nodos distribuidos en Ω . En (3.37) j representa los nodos sobre los cuales se coloca cada spline; $\beta_j(r)$ es el valor del spline asociado al nodo j , al evaluarlo en el voxel r perteneciente a su soporte; $\beta_j(r)$ está definido como en (A.6); α_{kj} es la contribución del spline j a la clase k .

Para encontrar todos los vectores de verosimilitud en el espacio de color, es necesario conocer la contribución de cada spline a cada clase k (α_{kj}). La solución de esta tarea es calculada al minimizar el funcional expresado por (3.38):

$$\min_{\alpha_k} \sum_{r \in D} \omega(r) \left(v_k(r) - \sum_j \alpha_{kj} \beta_j(r) \right)^2 + \tau \sum_{\langle m, n \rangle} (\alpha_{km} - \alpha_{kn})^2 = \varepsilon_k. \quad (3.38)$$

donde D representa el conjunto de voxels para los cuales se conoce su vector v ; $\omega(r)$ es una función que pesa la confiabilidad del vector v , obtenido durante el experimento de nombramiento del color. En este trabajo $\omega(r)$ se define como en (3.39):

$$\omega(r) = \begin{cases} a, & \text{hubo consenso.} \\ 1, & \text{en otro caso.} \end{cases} \quad (3.39)$$

En (3.39) se tiene que $a > 1$, y *consenso* indica que incluyen sólo aquellos voxels, cuyas muestras, durante el experimento psicofísico reciben una misma categoría por todos los sujetos. En (3.38) el segundo término se introduce para lograr suavidad entre los coeficientes de splines vecinos, lográndose así un modelo de tipo CAM para los coeficientes de los splines; τ es un parámetro positivo que controla el grado de suavidad.

Al calcular las derivadas respecto a cada coeficiente de los splines se tiene:

$$\frac{\partial \varepsilon_k}{\partial \alpha_{ki}} = -2 \sum_{r \in D} \omega(r) \left(v_k(r) - \sum_j \alpha_{kj} \beta_j(r) \right) \beta_i(r) + 2\tau \sum_{m \in |N_i|} (\alpha_{ki} - \alpha_{km}). \quad (3.40)$$

Al igualar a cero (3.40) se obtiene:

$$\frac{\partial \varepsilon_k}{\partial \alpha_{ki}} = - \sum_{r \in D} \omega(r) \left(v_k(r) - \sum_j \alpha_{kj} \beta_j(r) \right) \beta_i(r) + \tau \sum_{m \in |N_i|} (\alpha_{ki} - \alpha_{km}) = 0. \quad (3.41)$$

La notación $|N_i|$ representa la cardinalidad del conjunto de nodos vecinos para el spline i .

La expresión dada por (3.41) da lugar al sistema lineal de ecuaciones que a continuación se indica:

$$A\mathbf{b} = \mathbf{e}. \quad (3.42)$$

En (3.42) A , b y C se definen como a continuación:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix}. \quad (3.43)$$

En (3.43) los elementos de la diagonal (a_{ii}) están definidos como en (3.44) :

$$a_{ii} = \sum_{r \in D} \omega(r) \beta_i^2(r) + \tau |N_i|. \quad (3.44)$$

El resto de los elementos (a_{ij}) se expresan en (3.45):

$$a_{ij} = \sum_{r \in D} \omega(r) \beta_i(r) \beta_j(r) - I_{ij} \tau. \quad (3.45)$$

Los elementos de b se describen en (3.46)

$$\mathbf{b} = \begin{pmatrix} \alpha_{k1} \\ \alpha_{k2} \\ \dots \\ \alpha_{km} \end{pmatrix}. \quad (3.46)$$

Finalmente \mathbf{e} está dada por (3.47):

$$\mathbf{e} = \begin{pmatrix} \sum_{r \in D} \omega(r) p_k(r) \beta_1(r) \\ \sum_{r \in D} \omega(r) p_k(r) \beta_2(r) \\ \dots \\ \sum_{r \in D} \omega(r) p_k(r) \beta_m(r) \end{pmatrix}. \quad (3.47)$$

En (3.44) I_{ij} es una función indicadora que permite conocer si los nodos i e j son vecinos;

Una vez calculadas las contribuciones de cada spline a cada clase en cada nodo, es posible conocer v en todo el espacio de color al usar (3.37).

Conocido el campo v , la partición R del espacio se obtiene según (3.32). Con la vía de interpolación descrita se obtienen regiones conexas para cada categoría. En la sección (4) se podrán apreciar los resultados de este proceder, que fue el escogido como método de interpolación.

3.4. Método de segmentación propuesto

Como resultado de la modelación de categorías de color se obtiene un campo de verosimilitudes, que es interpretado como un diccionario probabilístico de categorías del color. Este diccionario será usado para la segmentación.

3.4.1. Segmentación lingüística del color a través de un modelo Bayesiano jerárquico

La fusión del modelo de categorización elaborado (ver sección 3.3.1) con técnicas Bayesianas de interacción espacial (ver sección (3.1)) es el tema al cual está dedicada esta sección y explica una nueva propuesta de segmentación en imágenes en colores.

Acorde a la nueva propuesta, cada región de color extraída en una imagen, tiene asignado un descriptor denotado como $\langle C, S \rangle$, donde C es el atributo de color y S es el atributo de intensidad. Los valores de ambos atributos se indican a continuación:

$C \in \{\text{rojo, azul, verde, amarillo, anaranjado, púrpura, gris, negro, blanco, rosa, café}\}$

$S \in \{\text{claro, oscuro}\}$

Estimación del atributo C

Para la estimación del atributo C , se calcula el campo de probabilidades p considerando los siguientes aspectos:

- Como verosimilitud ($P(D/C)$) se utiliza el diccionario probabilístico obtenido en el modelo de categorización elaborado. En la notación anterior D indica las observaciones o información de color de cada pixel r en la imagen.
- Para cada pixel r en la imagen se calcula un vector $p_r \in \mathbb{R}^n$ cuyas componentes ($p_k(r)$) expresan el grado de pertenencia del pixel r a cada una de las categorías de color usadas, es decir:

$$p_k(r) = P(C_r = k | D, I_n) \quad k = 1, \dots, n \quad (3.48)$$

En (3.48) I_n indica las interacciones entre las categorías básicas establecidas en [26] y que en este trabajo de tesis se definen según la matriz de interacciones dada en (3.30).

- Se exige suavidad del campo p a obtener.

Para la propuesta que se describe $n = 11$, ya que se consideran las 11 categorías del color dadas en [6].

En el cálculo de p se usa la generalización del modelo de campos de medida Markovianos Gaussianos con entropía controlada descrito en (3.3). Una vez conocido p , C_r se estima según (3.49):

$$C_r = \operatorname{argmax}_{k=\{1,2,\dots,n\}} p_{kr}. \quad (3.49)$$

En (3.49) se indica que el atributo de color C , asociado al pixel r , está en correspondencia con la componente de mayor valor en el vector $p(r)$ estimado según (3.28). Como se dijo anteriormente cada componente $p_k(r)$ del vector $p(r)$ expresa el grado de pertenencia del pixel r a la categoría k . El orden en que aparecen las componentes se corresponde con el orden en el que aparecen los valores que puede tomar el atributo C .

Para la estimación del atributo C se debe tomar en cuenta la incertidumbre presente en zonas oscuras de color. Con este fin, las componentes de cada vector de verosimilitudes (v) resultado de la modelación de las categorías, y correspondiente al contenido de color del pixel r (v_r), se ordenan como se indica en (3.50):

$$v_{k_1} \geq v_{k_2} \geq \dots \geq v_{k_n}. \quad (3.50)$$

En (3.50) $n=11$. Si se cumple que:

$$v_{k_1} < v_{k_2} * Thr. \quad (3.51)$$

cuando v_{k_1} es el contenido de la componente correspondiente a la categoría negra y v_{k_2} el correspondiente a cualquiera de las categorías cromáticas; entonces se enfatiza la componente cromática, intercambiando los valores de las componentes negra y cromática. Esto conlleva a que, durante el cálculo de p_r se realce el peso de las componentes de color en zonas muy oscuras. Thr es un umbral y en este trabajo se considera $Thr = 2$.

Estimación del atributo S

Luego de estimado el atributo C para cada pixel en la imagen, se procede a la subdivisión de cada categoría encontrada en dos subclases: *claro* y *oscuro*, este paso constituye la estimación del atributo S dado el atributo C y se le ha denominado etapa de refinamiento por categorías. En este paso la estimación de S se realiza a través de la minimización de (3.19), ya que en calidad de matriz de interacciones entre las componentes del vector p se usará la matriz identidad, pues no se consideran las interacciones entre colores y se necesita solamente subdividir cada una de las categorías. Para el cálculo del atributo S , en (3.17) la verosimilitud se halla como se expresa en (3.52) :

$$v_k(r) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(L^*(r)-\mu_k)^2}{2\sigma_k^2}}. \quad (3.52)$$

En (3.52) $k \in \{0, 1\}$ indica las subclases o modelos de intensidad *claro* y *oscuro*; $L^*(r)$ expresa el nivel de intensidad asociado al pixel r ; μ_k es el valor típico de intensidad para la subclase k ; los valores iniciales de los modelos μ_k se toman a partir del mínimo y máximo nivel de

intensidad de la categoría a subdividir. En (3.52) σ_k es la desviación estándar de los niveles de grises para cada modelo. Este parámetro será considerado constante, de modo que solamente se estimará la media de cada modelo.

La estimación de S se realiza usando un algoritmo de dos pasos nuevamente, en el primer paso se calcula p dados μ_k , en el segundo paso, dado p se calculan los nuevos μ_k . El paso de estimación de p se realiza según (3.20). Las componentes del vector p calculado para cada pixel r , expresan el grado de pertenencia al modelo claro y al modelo oscuro de intensidad.

El paso de estimación de μ_k se realiza según (3.53):

$$\frac{\partial L(p, \mu_k)}{\partial \mu_k} = 0 \Rightarrow \mu_k = \frac{\sum_{r \in L} p_k^2 L^*(r)}{\sum_{r \in L} p_k^2(r)}. \quad (3.53)$$

Esto se repite hasta llegar al criterio de convergencia que se tomó: $|(p(t+1) - p(t))| < \epsilon$, en nuestro caso $\epsilon = 10^{-4}$; t indica la iteración. Este criterio se verifica durante la subdivisión de cada categoría de color encontrada. Finalmente el atributo S se calcula como en (3.54):

$$S_r = \operatorname{argmax}_{k=\{1,2\}} p_{kr}. \quad (3.54)$$

Una vez terminada la estimación de los descriptores $\langle C, S \rangle$ se procede a una etapa de postprocesamiento. En esta etapa las regiones de color (Reg) cuya área en pixeles $A_{Reg} < \varepsilon$ son eliminadas.

Aunque la etapa de refinamiento descrita estuvo restringida a la subdivisión de cada categoría encontrada a dos subclases, la extensión a más es posible, lo que se podrá apreciar en (5.2).

La estimación del atributo C y seguidamente la estimación de S dado C , a través de la fusión de la estimación Bayesiana con un modelo de categorización del color, explican el nombre de la propuesta: ***Segmentación lingüística de imágenes en colores a través de un modelo Bayesiano jerárquico.***

En la Figura 3.3 se ilustra gráficamente la propuesta descrita.

A continuación se describen los experimentos realizados y resultados obtenidos al aplicar la propuesta explicada.

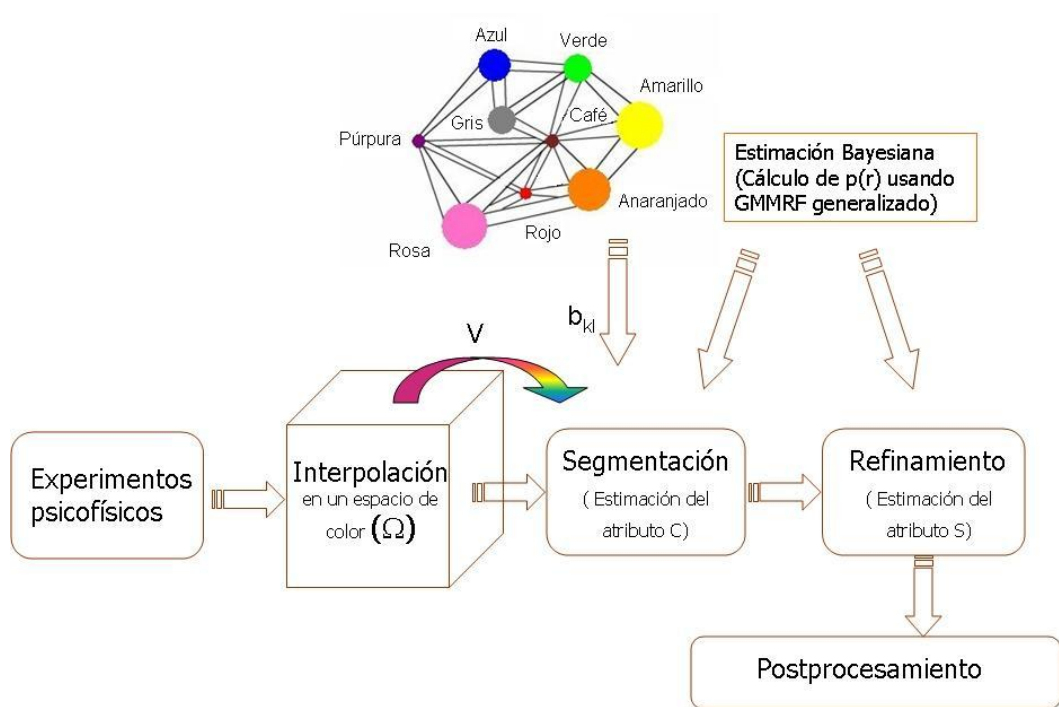


Figura 3.3. Representación gráfica de la propuesta de segmentación.

Capítulo 4

Trabajo realizado y discusión

En este capítulo se describen los experimentos hechos para el método de segmentación propuesto.

4.1. Construcción del conjunto de entrenamiento

Al usar las 11 categorías de color definidas por Berlin y Kay, un conjunto de 27 imágenes reales en colores y 32 observadores se hicieron dos experimentos de nombramiento del color. En el primero las zonas a nombrar tienen información de contexto, en el segundo no. Cada sujeto fue sometido a la prueba de Ishihara [61],[56] para detectar defectos en la visión del color.

Experimento con contexto

En cada una de las 27 imágenes se indicaron zonas, a las cuales el observador le asignó una de las 11 categorías básicas. En este experimento los participantes tuvieron no sólo los nombres de las 11 categorías básicas, sino también los prototipos de estos colores según [69] (vea Figura 4.1). La realización de este experimento permitió corroborar la influencia del contexto en el nombramiento de colores, sobre todo en aquellos casos donde la zona de color observada es descrita por más de una categoría. En términos de la teoría de información, lo dicho anteriormente quiere decir, que la entropía es alta y las respuestas entre sujetos ante una zona de color con tales características, es diferente. Más adelante se ilustra como, en las zonas descritas, la respuesta de un mismo observador es diferente, con y sin información contextual.

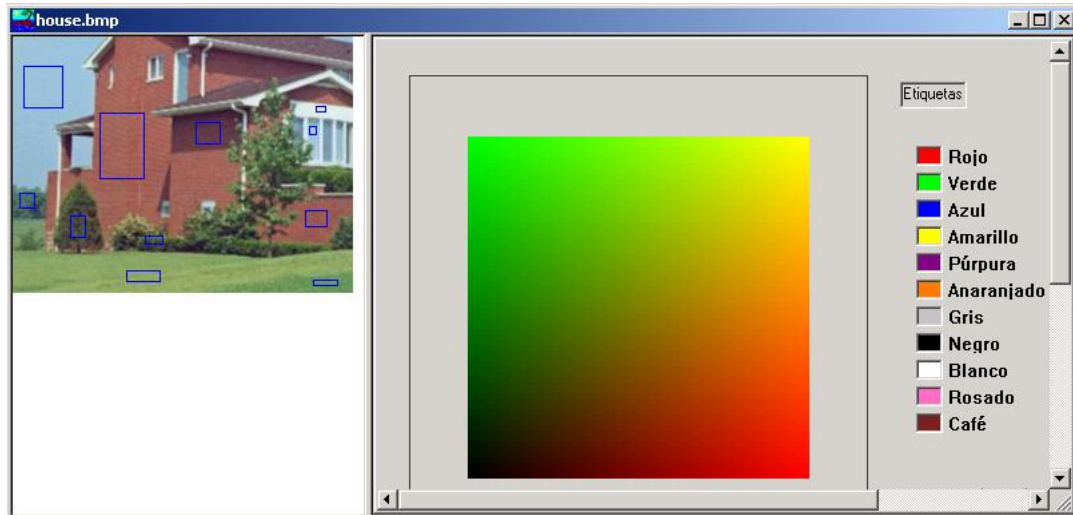


Figura 4.1. Experimento de nombramiento de color considerando contexto.

Experimento sin contexto

El número total de zonas de color señaladas en las imágenes de entrenamiento es de 336. Cuando el observador asigna un nombre de color a una región en la imagen que se le muestra, realiza una operación de promediación. Basado en esta acción perceptual, se calcularon los valores RGB promedio de cada región seleccionada en las 27 imágenes del experimento. Con estos valores se construyen muestras que se visualizan aleatoriamente a cada observador (vea Figura 4.2) y cada una de ellas aparece sobre un fondo gris. Entre la aparición de una y otra se visualiza un fondo gris para disminuir la influencia del efecto de *memoria de imagen* (*afterimage*, en inglés). En este experimento, el observador conoce solamente los nombres de los colores básicos a usar (vea Figura 4.2).

Comparación de los experimentos

La respuesta de un mismo observador en ambos experimentos es la misma cuando se está en presencia de una muestra de color que se describe sin duda alguna, con una sola categoría. En caso contrario la respuesta es diferente. En el primer caso la ubicación del color en un espacio permite asignar sin dudas, en la percepción, una de las categorías básicas; en el otro el color está en una zona de transición y resulta muy difícil asignar una sola categoría. En la Figura 4.3 se observa un ejemplo de lo dicho anteriormente para un mismo observador. En el

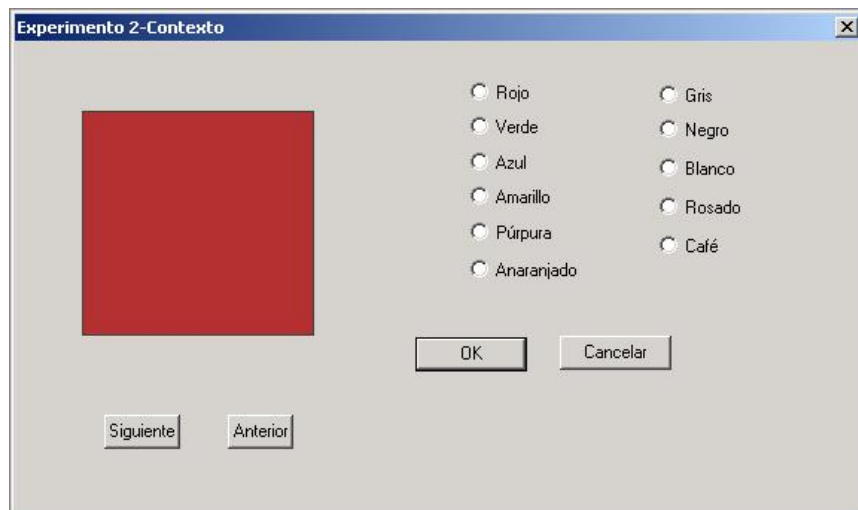


Figura 4.2. Experimento de nombramiento de color sin contexto.

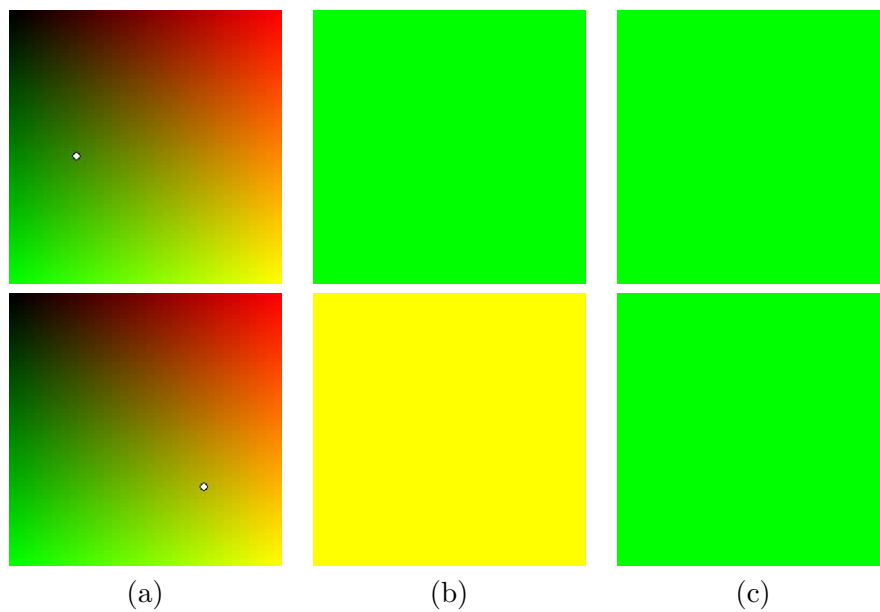


Figura 4.3. (a) Ubicación de una muestra de color en el espacio RGB, (b) respuesta de un observador en el experimento con contexto, (c) respuesta del observador en el experimento sin contexto.

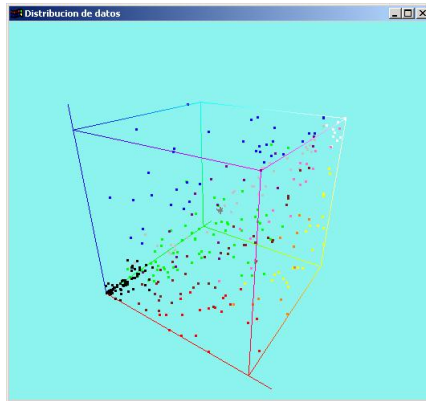
caso de muestras de color ubicadas en zonas de transición en el ejercicio de nombramiento del color influye notablemente el contexto. Este aspecto, junto al hecho de hacer independiente el experimento de nombramiento del color del conjunto de imágenes y los prototipos seleccionados, explican la razón por la que se consideraron los resultados del experimento sin contexto.

Después del experimento de nombramiento de color, cada muestra se describe a través de un vector cuyas componentes se calculan según (3.31). Este vector es denominado vector de verosimilitudes (ver sección 3.3.1).

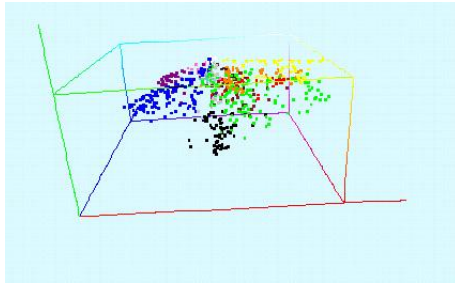
4.2. Modelación de las categorías de color en un espacio a través de la estimación de la verosimilitud

Con los vectores de verosimilitudes de cada muestra de color, se realiza el proceso de interpolación usando splines cuadráticos (sección 3.3.1). Esta vía de interpolación fue adecuada para los datos disponibles y la distribución de los mismos; mas adelante se muestran los resultados.

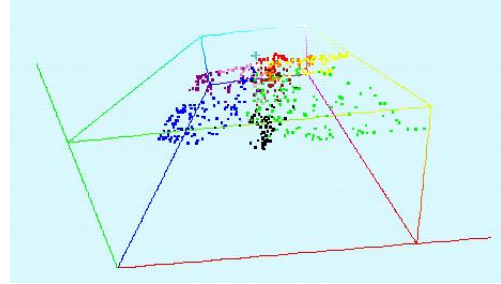
En este trabajo se usaron las representaciones de color RGB, $L^*a^*b^*$ y $L^*u^*v^*$. En la Figura 4.4 se ilustra la distribución de los datos experimentales, en dichos espacios. La etiqueta de color asignada a los voxels que se observan, está en correspondencia con la componente de mayor valor en el vector de verosimilitudes calculado experimentalmente (expresión 3.31).



(a)



(b)



(c)

Figura 4.4. Distribución de las muestras en los espacios (a) RGB, (b) $L^*a^*b^*$ y (c) $L^*u^*v^*$.

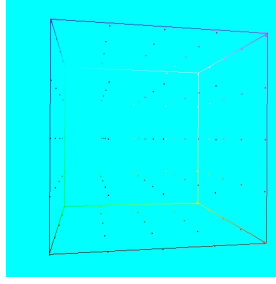


Figura 4.5. Distribución de los nodos para los splines cuadráticos en el espacio RGB usando $\Delta = 64$.

A continuación se describe el proceso de interpolación por splines cuadráticos que da lugar a la modelación de las categorías básicas. A través de este proceso se conoce la etiqueta de color más verosímil para cada voxel y la partición obtenida permite saber la forma geométrica aproximada de las categorías básicas de color.

4.2.1. Modelación en el espacio RGB

La distribución de los nodos sobre los cuales se colocan los splines cuadráticos aparece en la Figura 4.5. El espacio entre nodo y nodo es denotado como Δ y los valores usados fueron 16, 32 y 64.

En la Figura 4.6 se observa el comportamiento de la interpolación usando estos tres diferentes valores. Al comparar los resultados que se muestran en la Figura 4.6 puede observarse que $\Delta = 64$ es el valor para el cual se logró conectividad de las regiones de color. Este comportamiento fue el mismo para el resto de las otras categorías. Por otra parte se reduce el costo computacional en cuanto a tiempo, de horas (usando $\Delta = 16$) a 1.26 segundos al usar $\Delta = 64$.

Un parámetro muy importante es el que controla la suavidad entre las contribuciones de los splines (3.38). En este trabajo se experimentó con $\tau = \{0, 0.0001, 0.001, 0.01, 0.1\}$. Considerar valores muy pequeños ($\tau = 0.0001$, $\tau = 0$) conlleva a una aproximación muy local de los datos, dando lugar a regiones no suaves. Asimismo se obtienen zonas de color disjuntas, cuando estas corresponden a una misma categoría (ver Figuras 4.7 y Figuras 4.8(a) y (b)). Las discontinuidades en una misma categoría de color también se observan para valores de τ grandes ($\tau = 0.1$), obsérvese lo dicho anteriormente en la Figura 4.8(c). Los valores óptimos de τ , para la experimentación realizada, fueron encontrados entre 0.001 y 0.01. Partiendo de

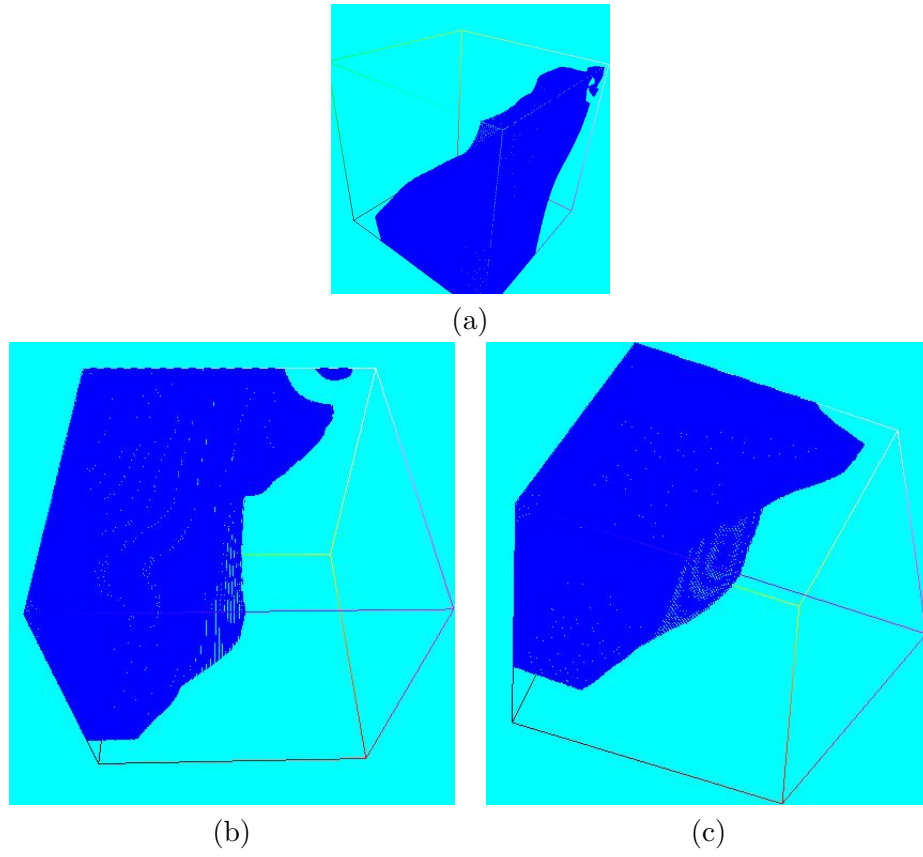


Figura 4.6. Resultados de la interpolación en el espacio RGB para la categoría azul usando $\tau = 0,01$ con (a) $\Delta = 16$, (b) $\Delta = 32$ y (c) $\Delta = 64$.

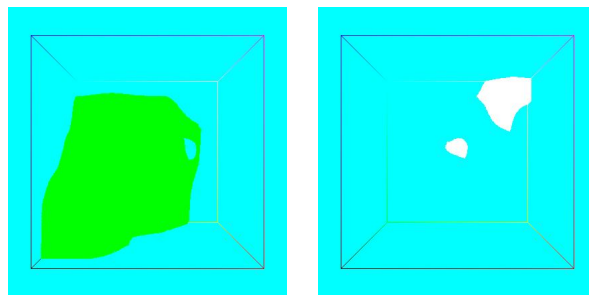


Figura 4.7. Resultados de la interpolación por splines cuadráticos usando $\tau = 0$.

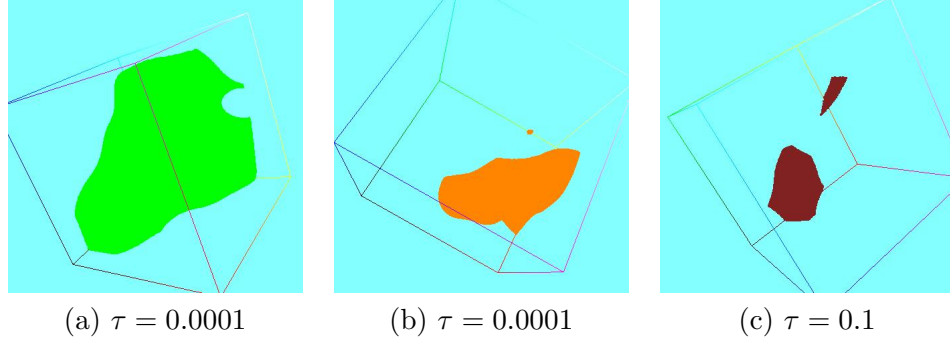


Figura 4.8. Zonas de color asignadas a las categorías: (a) verde y (b) anaranjado ; (c) zona asignada a la categoría café.

los resultados experimentales $\tau = 0.01$ fue seleccionado como adecuado.

Conocidos los α_{kj} se obtiene la partición que se ilustra en la Figura 4.9. Los niveles de grises (categorías acromáticas), aparecen a lo largo de la diagonal principal, mientras que las categorías cromáticas se distribuyen en el resto del espacio.

A partir de la expresión (3.31) se puede considerar que el vector de verosimilitudes calculado es una aproximación de una distribución de probabilidades $P(r)$ para cada punto del espacio. Usando $P(r)$ se calcula la entropía según (4.1):

$$Entropia(r) = - \sum_{k=1}^{11} P_k(r) \log(P_k(r)) \quad (4.1)$$

A través de este cálculo en la Figura 4.10 se ilustra el comportamiento de la incertidumbre en un corte dado en el espacio RGB. Mientras mayor es la incertidumbre (mayor entropía), más intenso es el nivel de gris (Figura 4.10(c)). Esto es característico en las zonas de transición entre categorías. Por otra parte las zonas de baja incertidumbre (zonas más oscuras) se corresponden con regiones donde hay menor duda al asignar una categoría (baja entropía). En la Figura 4.10(d) se ilustra la componente mayor de cada vector de verosimilitud encontrado en el corte representado. En ella los niveles de gris oscuros se corresponden con las componentes máximas más pequeñas (regiones de transición) y los niveles de gris más intensos se distribuyen en las zonas de color con menor entropía.

La elección de a en $\omega(r)$ (3.38, 3.39) se realizó teniendo en cuenta que el error entre el vector de verosimilitudes obtenido por la interpolación y el experimental para las muestras de consenso fuese pequeño. Esto se logró con $a = 5$.

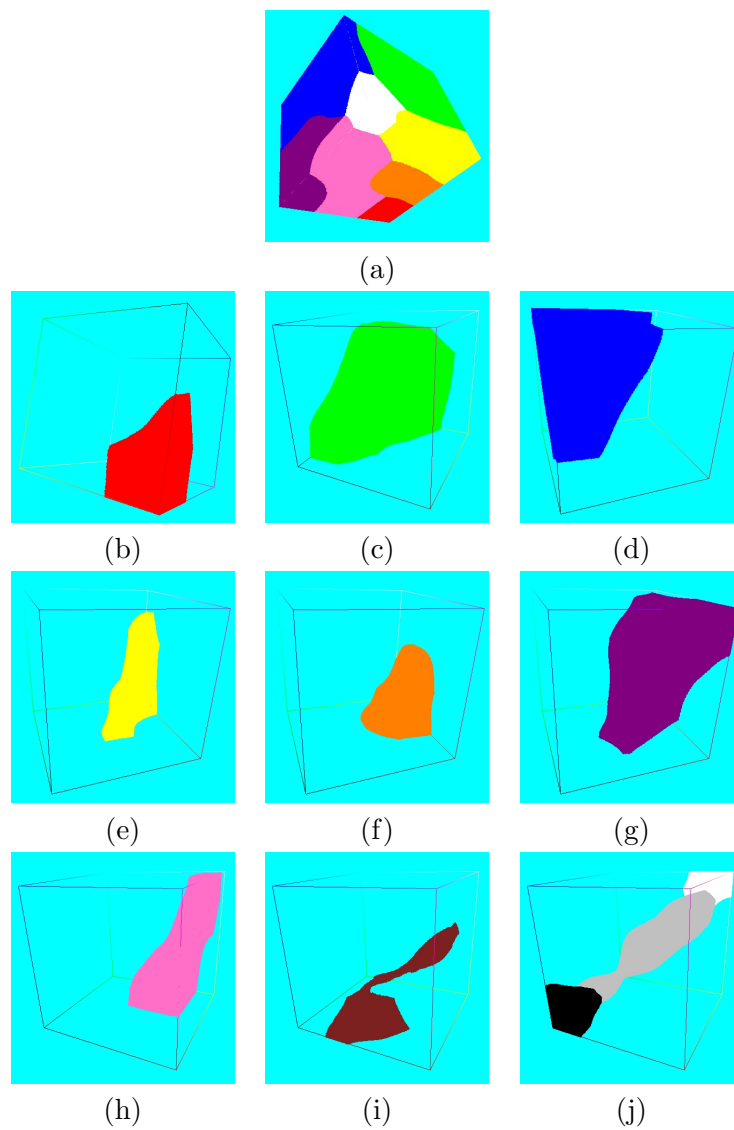


Figura 4.9. (a) Partición del cubo RGB en 11 clases. Localización de las categorías (b) rojo, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) café y (j) negro , gris y blanco en el cubo RGB.

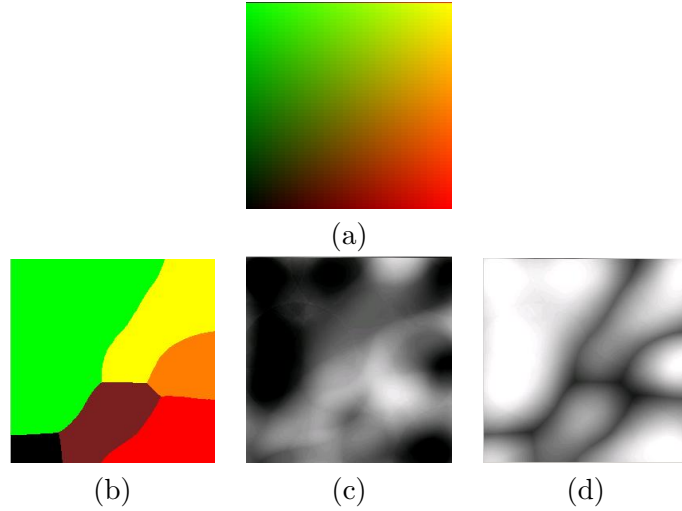


Figura 4.10. (a) Plano $B = 0$ en el espacio RGB, (b) partición obtenida, (c) representación de la entropía, (d) representación de la componente de mayor valor en el vector de verosimilitudes.

4.2.2. Modelación en los espacios $L^*a^*b^*$ y $L^*u^*v^*$

La ubicación de los nodos en los espacios CIE se puede realizar usando dos vías:

1. Construcción de retícula regular en los espacios $L^*a^*b^*$ y $L^*u^*v^*$, sobre la cual se distribuyen los nodos.
2. Transformación de las coordenadas de los nodos en RGB a coordenadas en los espacios $L^*a^*b^*$ y $L^*u^*v^*$, dando lugar a una retícula no regular de nodos.

A continuación se explica el desarrollo de las dos vías; posteriormente se comparan y finalmente se selecciona una de ellas como adecuada, teniendo en cuenta la calidad de la partición.

Construcción de retícula regular en los espacios $L^*a^*b^*$ y $L^*u^*v^*$

Teniendo en cuenta el rango de cada dimensión en los espacios $L^*a^*b^*$ y $L^*u^*v^*$, se construye la retícula regular. Para claridad en la lectura de los gráficos ilustrados, se indica al lector que el eje de color azul representa a L^* , el eje de color verde al v^* (b^*) y el eje rojo al u^* (a^*). El número de divisiones, así como el valor de τ óptimo, se determinó teniendo en cuenta la suavidad de las regiones de color obtenidas. En el caso del espacio $L^*u^*v^*$, los mejores resultados se obtuvieron usando una retícula de $10 \times 10 \times 10$, o sea 10 divisiones en cada una de las direcciones

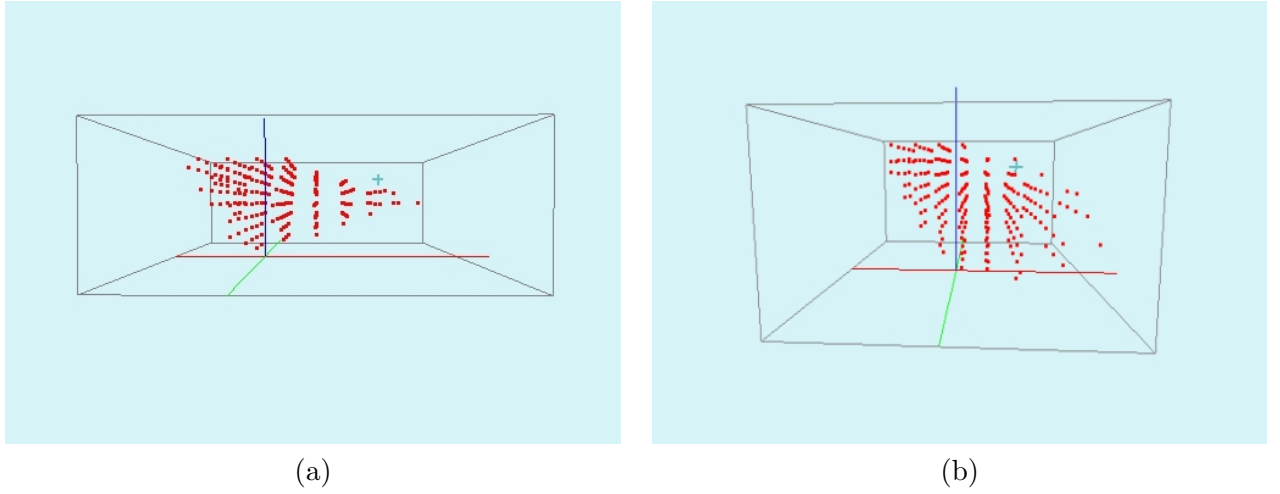


Figura 4.11. (a) Distribución de los nodos en el espacio $L^*u^*v^*$, (b) distribución de los nodos en el espacio $L^*a^*b^*$. Reticulado regular.

del espacio. Para el espacio $L^*a^*b^*$ la retícula fue $5 \times 5 \times 10$. En la Figura 4.11 se ilustra la distribución de los nodos para el reticulado hecho, considerándose sólo aquellos nodos cuya transformación de coordenadas $L^*a^*b^*$ ($L^*u^*v^*$) a RGB existe. En ambos espacios el mejor valor de τ fue 0.1. En las Figuras 4.12 y 4.13 se observan las particiones obtenidas. Obsérvese que para el espacio $L^*a^*b^*$ no se logró continuidad en el eje de grises (Figura 4.13(j)). Otras retículas o valores de τ experimentados conllevan a discontinuidades de las regiones de color pertenecientes a una misma categoría o a la desaparición de alguna clase de color en la partición. En las Figuras 4.14 y 4.15 se ejemplifica lo dicho anteriormente. Observe discontinuidades de la partición $L^*u^*v^*$ en regiones asociadas a una misma categoría (Figura 4.14 (a), (b), (c), (d)) o desaparición de zonas de color relativas a una categoría (voxeles asociados a la categoría blanca, Figura 4.14(e)).

En los experimentos mostrados en la Figura 4.15 (espacio $L^*a^*b^*$), se observan zonas de color relativas a una misma categoría que son disjuntas (Figuras 4.15(a), (b), (c) y (d)). El resultado dado en Figuras 4.15 (b) ilustra además la discontinuidad obtenida en el eje de grises y la desaparición de la categoría blanca en la zona de mayor valor del nivel de intensidad (eje L^*).

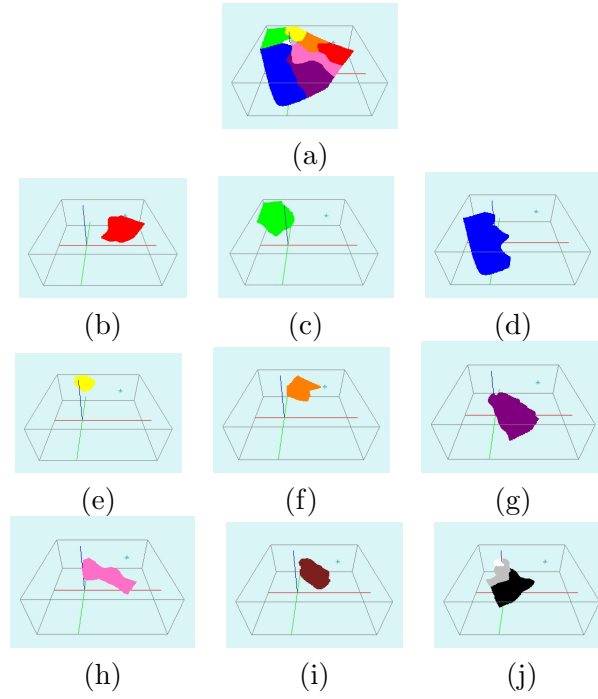


Figura 4.12. (a) Partición del espacio $L^*u^*v^*$ en 11 clases usando retícula regular. Localización de las categorías (b) roja, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) café y (j) negro , gris y blanco.

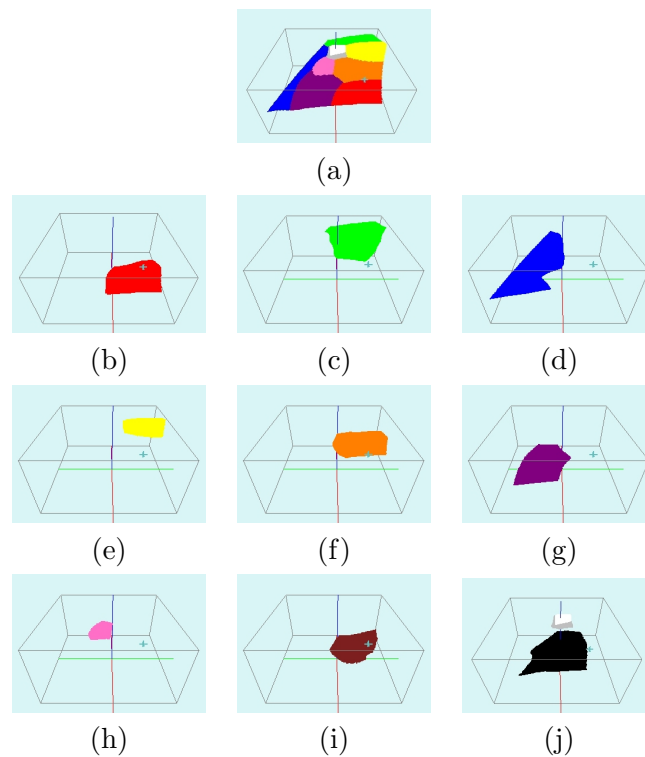


Figura 4.13. (a) Partición del espacio $L^*a^*b^*$ en 11 clases usando retícula regular. Localización de las categorías (b) roja, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) café y (j) negro , gris y blanco.

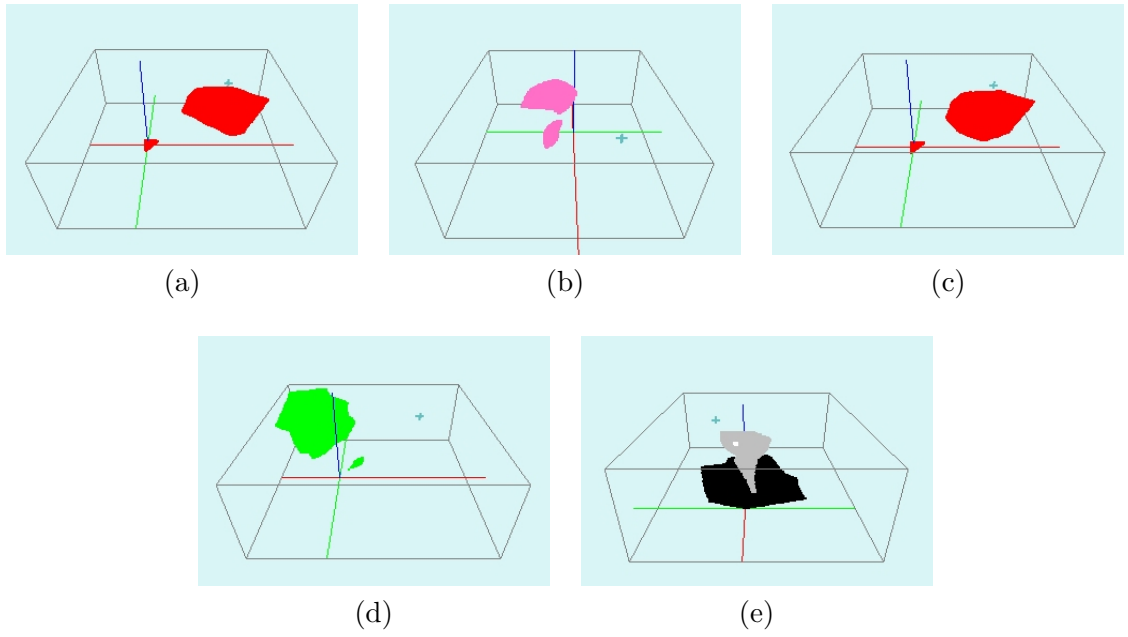


Figura 4.14. Otros experimentos en el espacio $L^*u^*v^*$, (a) $\tau = 0.01$, retícula de $5 \times 5 \times 10$, (b) $\tau = 0.1$, retícula de $5 \times 5 \times 10$, (c) $\tau = 1$, retícula de $5 \times 5 \times 10$, (d) $\tau = 0.01$, retícula de $10 \times 10 \times 10$, (e) $\tau = 1$, retícula de $10 \times 10 \times 10$.

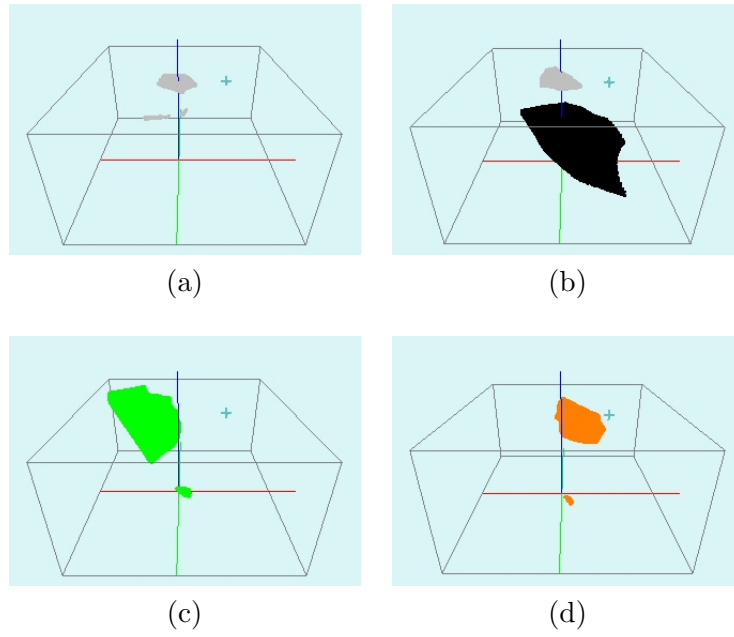


Figura 4.15. Otros experimentos en el espacio L^*a^*b , (a) $\tau = 0.01$, retícula de 5x5x10, (b) $\tau = 0.1$, retícula de 5x5x10, (c) $\tau = 0.01$, retícula de 10x10x10, (d) $\tau = 1$, retícula de 10x10x10.

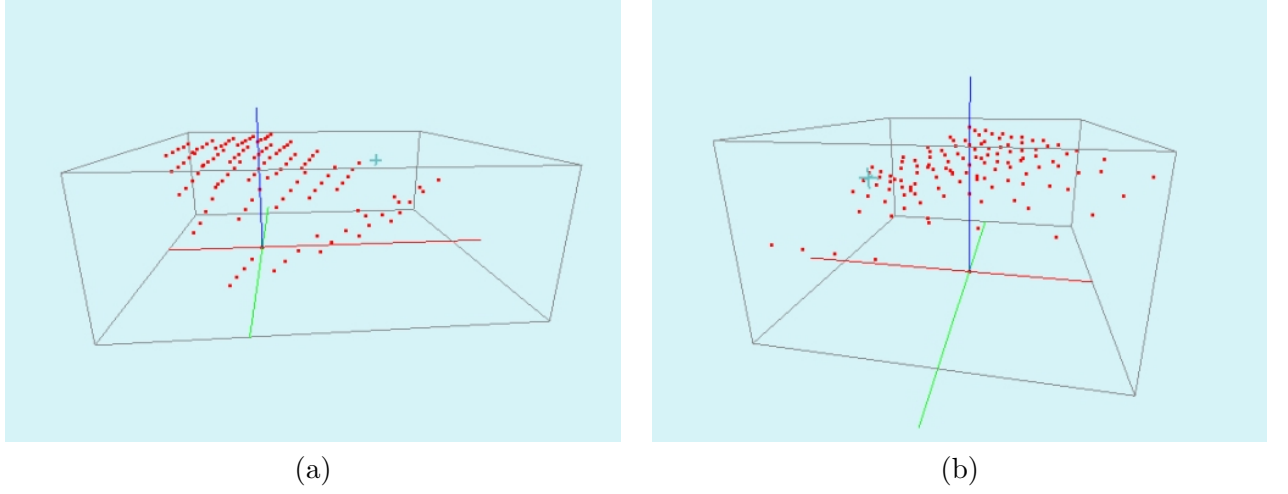


Figura 4.16. (a) Distribución de los nodos en el espacio $L^*u^*v^*$, (b) Distribución de los nodos en el espacio $L^*a^*b^*$. Los nodos son resultado de llevar las coordenadas de los nodos RGB a coordenadas $L^*u^*v^*$ y $L^*a^*b^*$.

Transformación de las coordenadas de los nodos en RGB a coordenadas en los espacios $L^*a^*b^*$ y $L^*u^*v^*$

Luego de la transformación de las coordenadas de los nodos en RGB a coordenadas en los espacios $L^*a^*b^*$ y $L^*u^*v^*$ usando las expresiones (2.7), (2.8), (2.9), (2.10), (2.11), (2.12) y (2.13), se obtiene la representación que se ilustra en la Figura 4.16.

Debido a que la retícula de nodos obtenida no es regular, surge la dificultad de la elección de Δ , o sea ¿cómo elegir el espaciamiento entre los nodos ubicados en la retícula irregular?. En este trabajo se ha calculado la distancia de cada nodo (i) a todos sus vecinos (m); y como Δ para ese nodo se considera la mayor de todas las distancias ($Dist$) calculadas. La expresión (4.2) describe lo planteado.

$$\Delta_i = \max_{m \in N_i} Dist(i, m). \quad (4.2)$$

En (4.2) N_i denota los vecinos del nodo i . Los mismos se obtienen a partir del sistema de vecindad del nodo RGB correspondiente, ubicado en una retícula que si es regular. Como resultado de la transformación del espacio RGB a los espacios $L^*a^*b^*$ y $L^*u^*v^*$ (sección 2.1.5) el volumen dado en el espacio RGB es distorsionado; puntos cercanos en un espacio pueden

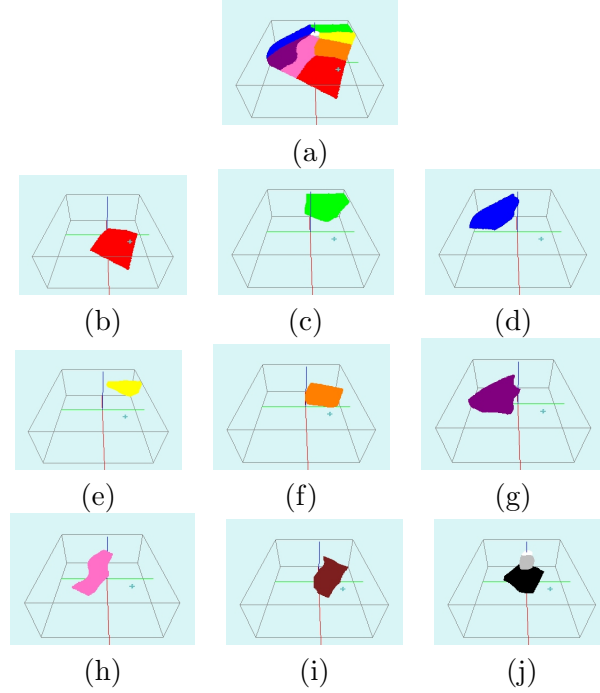


Figura 4.17. (a) Partición del espacio $L^*u^*v^*$ en 11 clases usando transformación de coordenadas de nodos RGB a nodos $L^*u^*v^*$. Localización de las categorías (b) roja, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) café y (j) negro , gris y blanco.

alejarse en el otro y viceversa. Con el criterio de elección dado en (4.2) se garantiza al menos, que el soporte de los splines obtenido en el espacio RGB esté contenido en el soporte de los espacios $L^*u^*v^*$ y $L^*a^*b^*$ y así respetar la distribución de las muestras usadas en el experimento de nombramiento de color.

El resultado de la interpolación aparece en las Figuras 4.17 y 4.18.

Comparación de las vías de reticulado

En el resultado obtenido a través del reticulado irregular hay una mejor partición. Esta afirmación se basa en las siguientes observaciones:

1. Mejor clasificación en las categorías acromáticas (compare Figuras 4.12(j) y 4.17(j) , y Figuras 4.13(j) y 4.18 (j)).
2. El vector de verosimilitudes obtenido se corresponde con la descripción de zonas de color en un espacio.

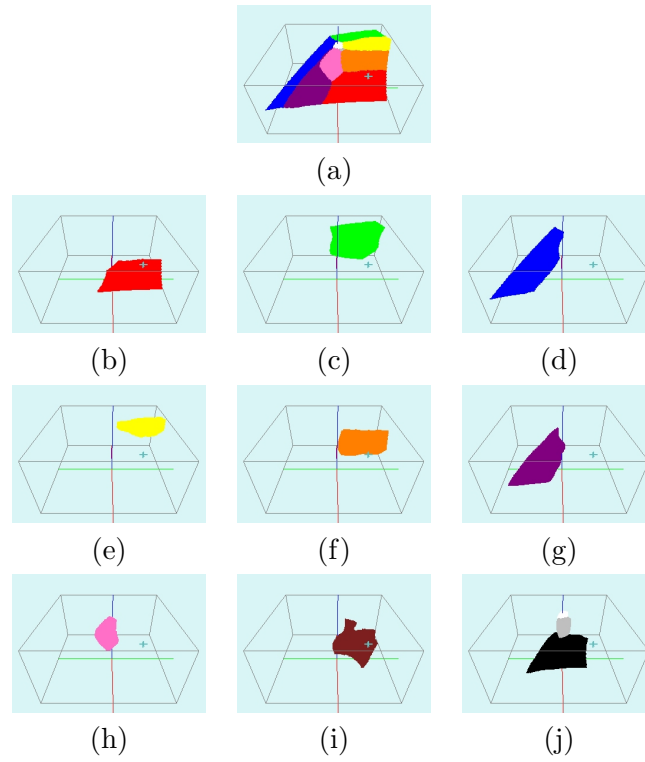


Figura 4.18. (a) Partición del espacio $L^*a^*b^*$ en 11 clases usando transformación de coordenadas de nodos RGB a nodos $L^*a^*b^*$. Localización de las categorías (b) roja, (c) verde, (d) azul, (e) amarillo, (f) anaranjado, (g) púrpura, (h) rosa, (i) café y (j) negro , gris y blanco.

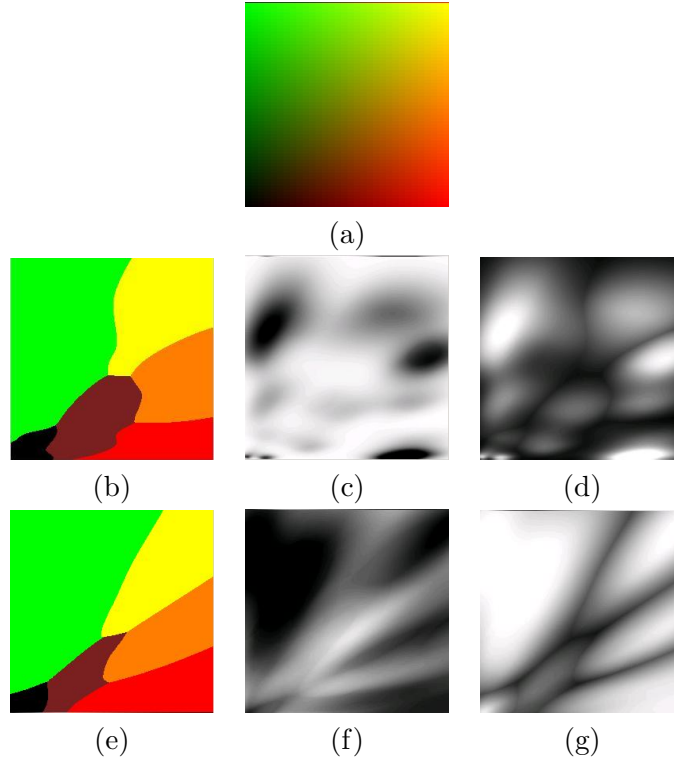


Figura 4.19. (a) Plano $B=0$ en el espacio RGB, (b) partición del espacio $L^*u^*v^*$ con retícula regular, y representada en el espacio RGB (plano $B=0$), (c) mapa de entropía, (d) componente de mayor valor en el vector de probabilidades, (e) partición del espacio $L^*u^*v^*$, usando reticulado no regular, y representada en el espacio RGB (plano $B=0$), (f) mapa de entropía, (g) componente de mayor valor en el vector de probabilidades.

Con el objetivo de argumentar la segunda observación se analiza el comportamiento de la entropía según expresión (4.1) en uno de los espacios CIE. En la Figura 4.19 se ilustra una zona de las particiones logradas en el espacio $L^*u^*v^*$, usando reticulado regular y no regular, y es representada en el corte correspondiente del espacio RGB. En la ilustración aparece además, el mapa de entropía asociado y una representación de la componente de mayor valor en el vector de verosimilitudes asociado al voxel. En la Figuras 4.19(c) y (f) mientras más oscuro es el nivel de gris, menor es la entropía. En las Figuras 4.19(d) y (g), mientras más intenso es el nivel de gris, mayor es el valor de la componente del vector de verosimilitudes que se ilustra.

El comportamiento dado en la Figura 4.19 es similar para otros cortes del espacio RGB. Al analizar la entropía en ambos reticulados, se concluye que el reticulado no regular permite simular mucho mejor la incertidumbre en zonas donde es muy difícil describir un color, así como

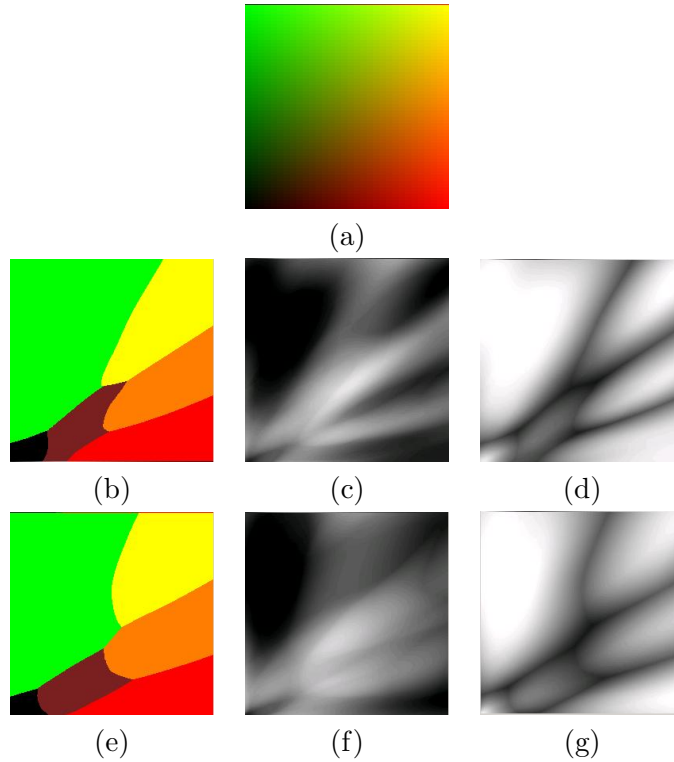


Figura 4.20. (a) Plano $B=0$ en el espacio RGB, (b) partición del espacio $L^*u^*v^*$ (reticulado no regular), representada en el espacio RGB (plano $B=0$), (c) mapa de entropía, (d) componente de mayor valor en el vector de probabilidades, (e) partición del espacio $L^*a^*b^*$ (reticulado no regular), representada en el espacio RGB (plano $B=0$), (f) mapa de entropía, (g) componente de mayor valor en el vector de probabilidades.

describir sin dudas el color, cuando este corresponde a una zona, donde el nombramiento del color ocurre con menos dificultad. Lo dicho anteriormente es muy importante pues define la calidad del diccionario probabilístico logrado y por lo expresado antes este diccionario es mejor al usar el reticulado no regular. La experimentación realizada indica que el reticulado no regular en los espacios CIE captura correctamente la distribución de los datos del experimento de nombramiento de color realizado.

Luego de seleccionar el reticulado no regular como adecuado se comparan los resultados entre los espacios RGB y CIE.

En la Figura 4.20 se observa nuevamente el comportamiento de la entropía en el espacio $L^*u^*v^*$ y se representa además lo que sucede en el el espacio $L^*a^*b^*$. Si se observa la Figura 4.10 y se compara el comportamiento de la entropía ilustrado en ella, con el ilustrado en las Figura 4.20 se puede decir que el espacio $L^*u^*v^*$ refleja mejor la incertidumbre cerca de las fronteras

CATEGORÍA	L*a*b*	L*u*v*	RGB
Rojo	4.33	4.14	3.91
Verde	28.21	28.92	29.72
Azul	23.25	22.74	22.74
Amarillo	8.52	7.76	6.96
Púrpura	14.48	13.918	14.49
Anaranjado	7.70	7.29	7.87
Gris	3.39	3.38	3.57
Negro	1.327	1.24	1.64
Blanco	0.97	0.92	1.03
Rosa	6.06	7.49	5.89
Café	1.71	2.16	2.13

Tabla 4.1. Porcentaje que ocupa cada categoría en los espacios de color RGB, L*u*v* y L*a*b*.

entre categorías y arroja una entropía menor en zonas donde perceptualmente está definida, la asignación a una categoría. Lo dicho anteriormente explica el porque del uso de los resultados de la partición L*u*v* en la etapa de segmentación.

A partir de las particiones en los espacios RGB, L*u*v* y L*a*b* usando reticulado no regular se calculó el porcentaje que ocupan, en todo el espacio, cada una de las categorías básicas de color. En la tabla 4.1 se detalla al respecto. Las categorías verde y azul reportan los mayores porcentajes, mientras que el blanco ocupa la menor parte del espacio.

4.3. Segmentación lingüística de imágenes en colores mediante un modelo Bayesiano jerárquico

A continuación se describen los experimentos que ilustran la propuesta de segmentación descrita. Para claridad en la lectura de los resultados, en la Figura B.1 se indica la leyenda empleada al codificar los atributos C y S.

4.3.1. Uso directo del diccionario probabilístico

En la sección 4.2 se ilustraron las particiones obtenidas en los espacios de color RGB, L*a*b* y L*u*v*. En cada una de ellas a cada voxel le fue asignada la etiqueta correspondiente a la componente de mayor valor en el vector de verosimilitudes calculado. En los ejemplos dados en las Figuras 4.21, 4.22, 4.23, 4.24 y 4.25 a cada pixel de la imagen original se le asigna la

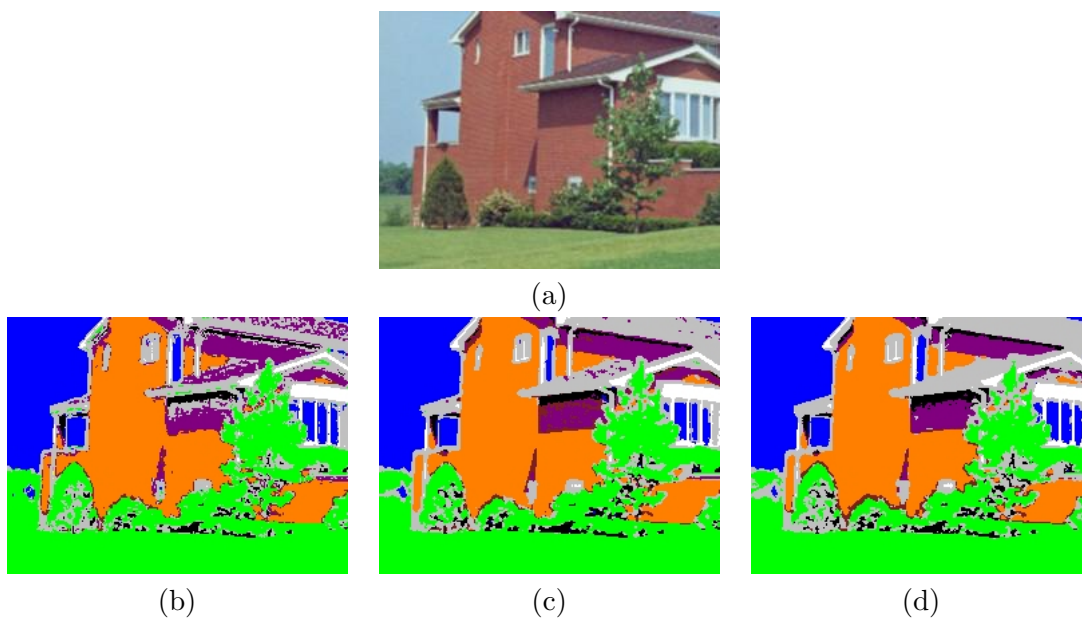


Figura 4.21. (a) Imagen original. Segmentación usando la partición en (b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$.



Figura 4.22. (a) Imagen original, segmentación usando la partición en (b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$.



(a)



(b)



(c)



(d)

Figura 4.23. (a) Imagen original, segmentación usando la partición en b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$.

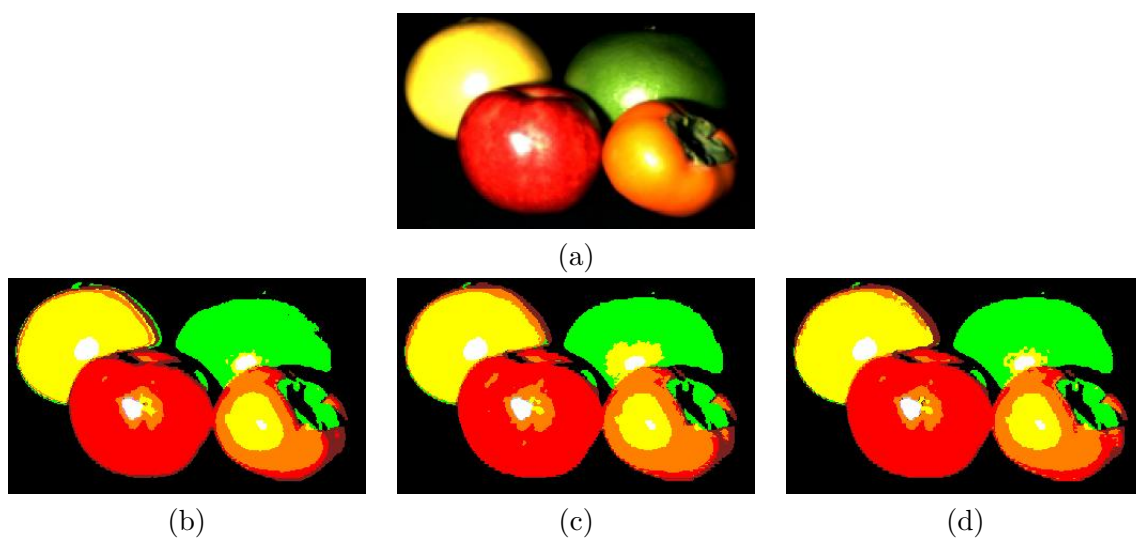


Figura 4.24. (a) Imagen original, segmentación usando la partición en (b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$.

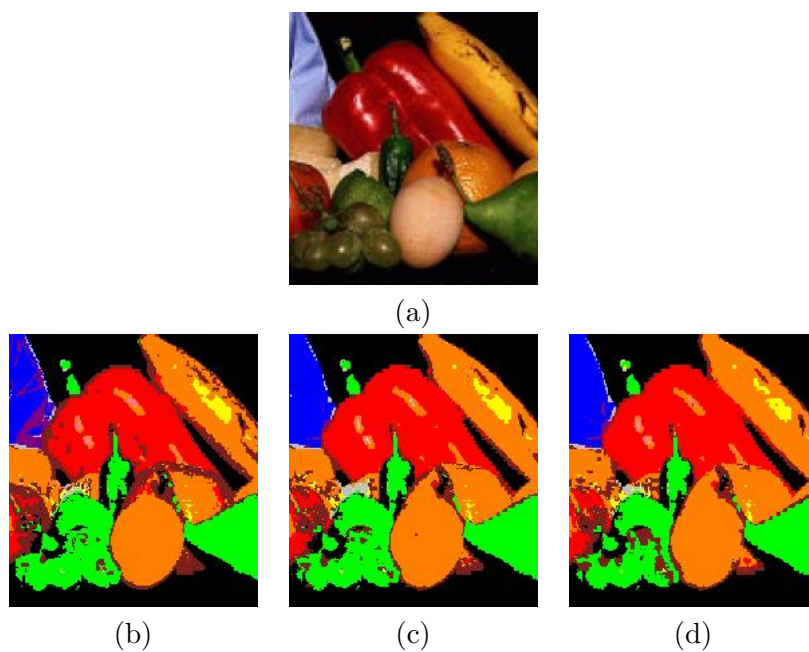


Figura 4.25. (a) Imagen original, segmentación usando la partición en (b) espacio RGB, (c) espacio $L^*u^*v^*$, (d) espacio $L^*a^*b^*$.



Figura 4.26. (a)Imagen original, (b) segmentación usando la partición en el espacio $L^*u^*v^*$.

etiqueta dada al voxel correspondiente a su contenido de color en RGB, $L^*a^*b^*$ y $L^*u^*v^*$. Con este proceder se hace uso directo del diccionario probabilístico encontrado en los espacios usados.

En general los resultados en los espacios $L^*u^*v^*$ y $L^*a^*b^*$ reflejan una mejor correspondencia con nuestra percepción. Teniendo en cuenta la experimentación realizada en la sección 4.2 y la investigación realizada en [22] y en [31] en este trabajo se ha seleccionado el espacio $L^*u^*v^*$ para la experimentación que en lo adelante se explica.

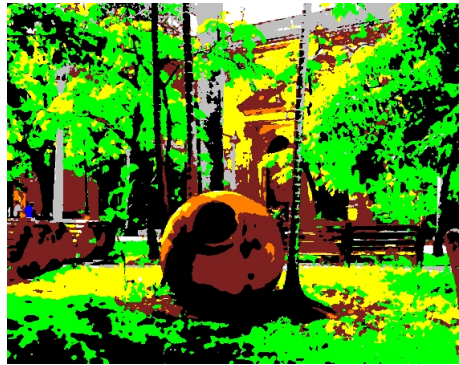
Usar el diccionario probabilístico directamente, resulta apropiado cuando se está en presencia de zonas muy homogéneas de color (ver Figuras 4.21, 4.22, 4.23, 4.24 y 4.25); sin embargo en las imágenes reales no sólo aparecen zonas homogéneas y en estos casos el enfoque Bayesiano junto a campos aleatorios Markovianos es de gran utilidad. En la Figura 4.26 se observa un ejemplo de una imagen real, rica en detalles. Una correspondencia directa entre el contenido de color de cada pixel y el voxel en el espacio de color da resultados con una alta granularidad. En la sección (4.3.2) se mostrará la segmentación de esta imagen, a través de la nueva propuesta.

4.3.2. Estimación del atributo C

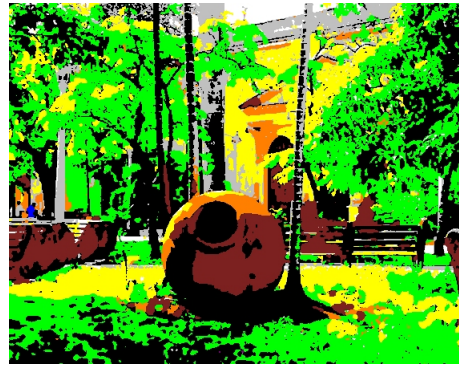
Usando el diccionario probabilístico obtenido luego de la estimación de la verosimilitud se realiza el cálculo del atributo C (3.4.1) a través de (3.24). Ejemplos de esta estimación aparecen en la Figura 4.27. Al comparar los resultados de aplicar (3.17) y (3.24) se puede observar que el uso de las interacciones entre colores da lugar a resultados con mayor correspondencia respecto al juicio humano. En la Figura 4.28 a) aparece seleccionada una zona del follaje de la Figura



(a)



(b)



(c)

Figura 4.27. (a) Imagen original, (b) segmentación obtenida al estimar C a través de (3.17), (c) segmentación obtenida al estimar C usando (3.24).

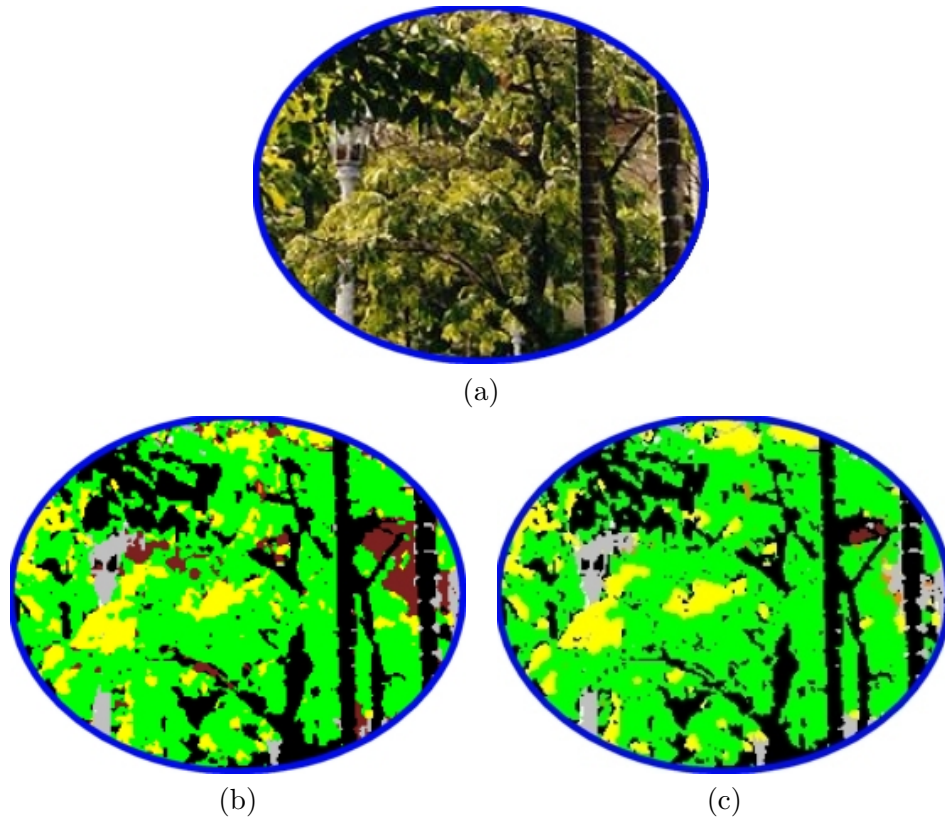
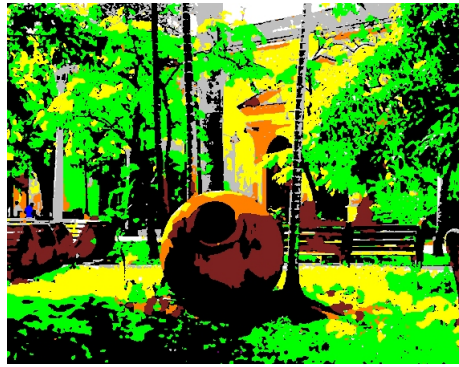


Figura 4.28. (a) Zona del follaje ampliada en imagen original de la Figura 4.28, (b) ampliación sin considerar interacciones (expresión (3.17)), (c) ampliación considerando interacciones (expresión (3.24)).

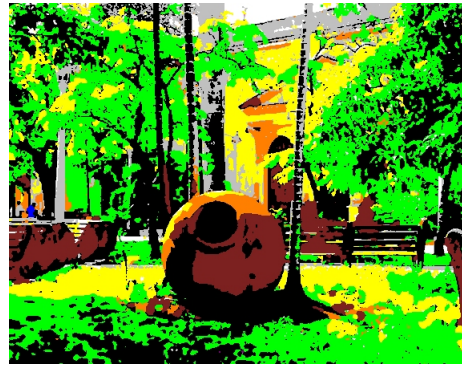
4.27(a). Si se comparan los resultados de usar la expresión (3.17) (Figura 4.28(b)) con los obtenidos al usar la expresión (3.24)(Figura 4.28(c)) se puede observar como la relación entre los colores gris, verde y café (vea zona del farol); y entre amarillo y verde (zona de follaje) se manifiestan, lo que se corresponde con las conclusiones dadas por Boynton y Olson en [26] y que fueron consideradas en la generalización presentada en esta tesis, del potencial propuesto en [30]. En el ejemplo mostrado en la Figura 4.27(c) se modeló la incertidumbre en zonas oscuras de color. En la Figura 4.29 se hace una comparación de este resultado con el que se obtiene cuando no se considera incertidumbre. Cuando no se considera la incertidumbre en zonas oscura de color, la estimación del atributo C en esas zonas conlleva a obtener etiquetas acromáticas (color negro en la puerta detrás del jarrón que se observa en la Figura 4.29). En la Figura 4.30 se ilustra otro ejemplo de la estimación del atributo C, con y sin considerar la incertidumbre (observe follaje al fondo).



(a)



(b)



(c)

Figura 4.29. (a) Imagen original, (b) segmentación obtenida al estimar C usando (3.24) sin considerar incertidumbre en zonas oscuras, (c) segmentación obtenida al estimar C usando (3.24) considerando incertidumbre en zonas oscuras.

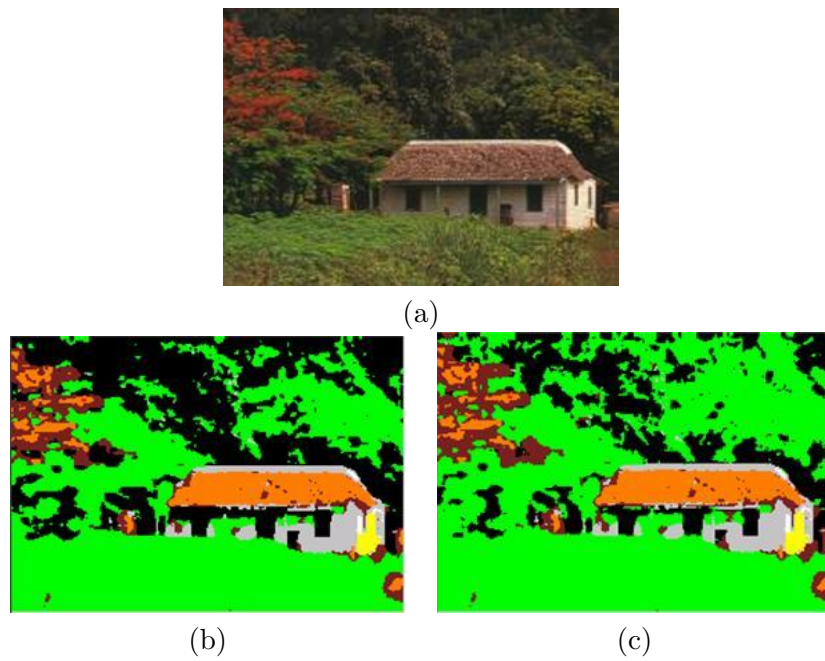


Figura 4.30. (a) Imagen original, (b) segmentación obtenida al estimar C usando (3.24), sin considerar incertidumbre en zonas oscuras de color, (c) segmentación obtenida al estimar C usando (3.24), considerando incertidumbre.

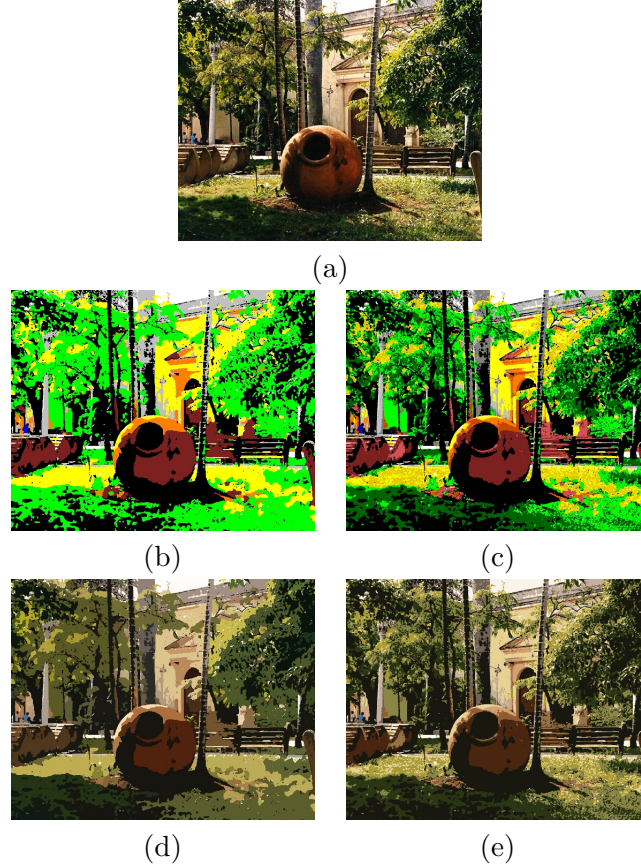


Figura 4.31. (a) Imagen original, (b) segmentación obtenida al estimar C usando (3.24), (c) segmentación obtenida al estimar S dada la estimación de C , (d) representación de la estimación de C usando la información de color en la imagen, (e) representación del atributo S usando la información de color en la imagen.

4.3.3. Estimación del atributo S

Luego de la estimación del atributo C , se estima el atributo S , con el objetivo de refinar cada categoría usando los atributos *claro* y *oscuro*. Este refinamiento por cada categoría (estimación del atributo C dado S) permite hacer una descripción más detallada de las regiones de color extraídas, especificando no solo una de las categorías básicas, sino también la característica relacionada con la intensidad. Más adelante se podrá apreciar el uso de esta parte de la propuesta para tareas de compresión y filtrado de ruido en imágenes en colores (ver sección 5).

En la Figura 4.31 se puede apreciar un ejemplo de estimar los atributos C y S . Las imágenes de las Figuras 4.31 (d) y (e) se obtienen luego de promediar el contenido RGB de todos los

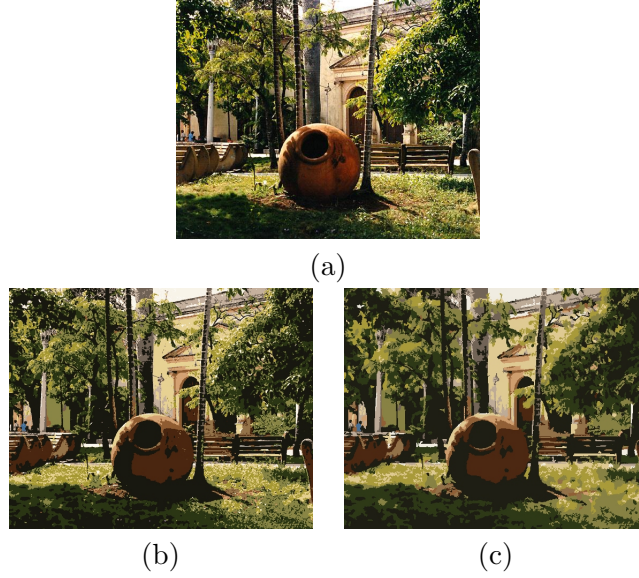


Figura 4.32. (a) Imagen original, (b) representación de la estimación de S dado C usando la información de color de imagen, (c) representación de la estimación de S dado C usando la información de color de imagen, luego de eliminar componentes conexas con área menor que cierto número de píxeles.

píxeles asociados a una etiqueta k , la cual está en correspondencia con la codificación usada al mostrar el descriptor $\langle C, S \rangle$. Para lograr la representación explicada se calcula el promedio por cada componente de color Ch_{ik} , donde $i \in \{R, G, B\}$, incluyendo sólo la información de color de los píxeles (r) asociados a etiqueta k . En la expresión (4.3) se describe lo planteado.

$$Ch_{ik} = \frac{\sum_{r \in L} \delta(I(r) - k) Canal_i(r)}{\sum_{r \in L} \delta(I(r) - k)}. \quad (4.3)$$

En (4.3), $Canal_i(r)$ denota la componente i de color en el píxel r ; $I(r) \in \{0, 1, \dots, E-1\}$ y denota la imagen de etiquetas, E es el número total de etiquetas. Otra posible forma de representación de la segmentación es aplicar la operación descrita en (4.3) sólo sobre las componentes conexas presentes en el resultado de la estimación de los atributos C y S . Para este caso, en la expresión (4.3) k denota una componente conexa. En la sección (5) se hará uso de la representación explicada.

Luego del refinamiento, se hace un postprocesamiento, con el objetivo de eliminar regiones muy pequeñas.

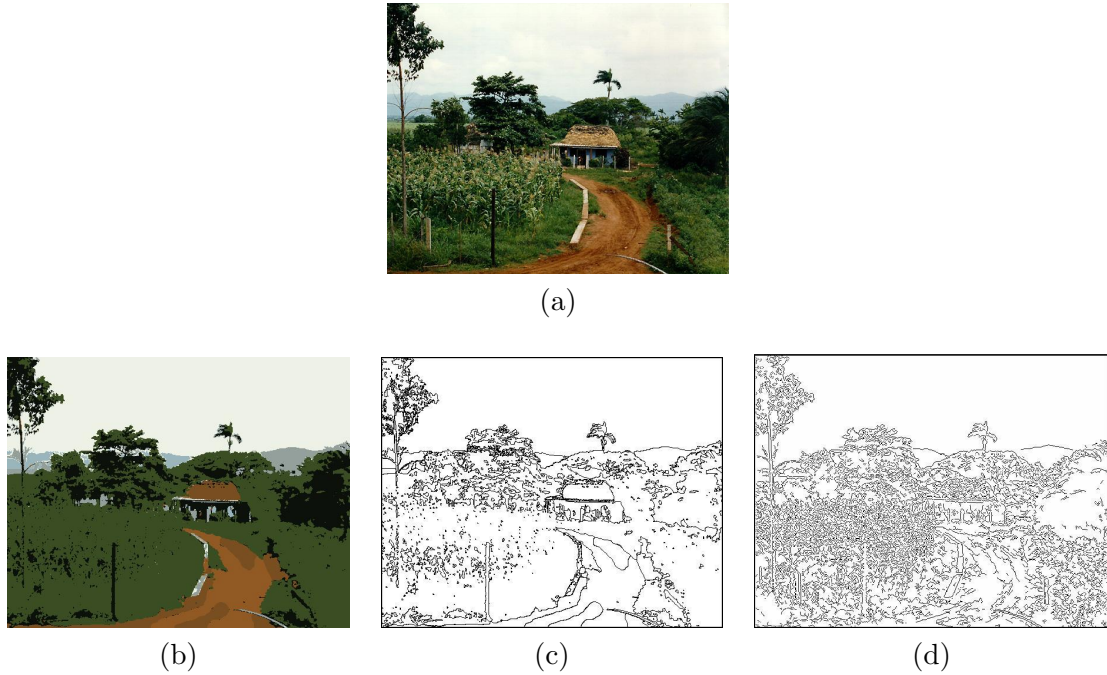


Figura 4.33. (a) Imagen original, (b) segmentación usando la nueva propuesta (solo se muestra representación del atributo C , usando expresión (4.3)), (c) bordes detectados a partir de la estimación del atributo C , (d) bordes detectados por Canny.

En la Figura 4.32(c) se muestra un ejemplo de postprocesamiento.

En el apéndice B se muestran otras imágenes segmentadas con el método propuesto. La elección de los valores adecuados de λ y μ se hizo experimentalmente considerando la correspondencia entre la segmentación estimada y el juicio humano. Según lo dicho anteriormente se tomó como valor adecuado $\lambda = 1$, $\mu = 0.5$. El rango de niveles de gris de todas las imágenes es normalizado entre 0 y 1, al realizar el refinamiento.

4.3.4. Comparación de la propuesta con otros métodos

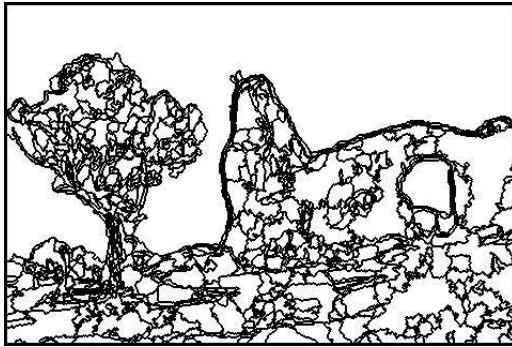
Como aspecto interesante de la propuesta de segmentación desarrollada es que permite detectar contornos cerrados. Los resultados se comparan con los obtenidos con el conocido detector de bordes de Canny [47].

En la Figura 4.33(c) se observan los contornos más significativos presentes en la imagen y sin discontinuidades a diferencia del resultado dado por Canny (Figura 4.33(d)).

A continuación se comparan resultados de algoritmos basados en detección de modos en un espacio de características [22], en gradiente de color y gradiente de texturas [70], [71] con la propuesta descrita. En las Figuras 4.34 y 4.35 se ilustran las comparaciones.



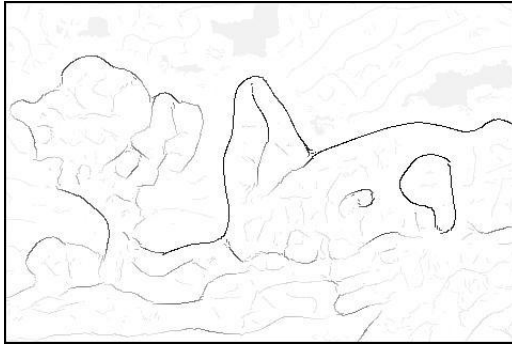
(a)



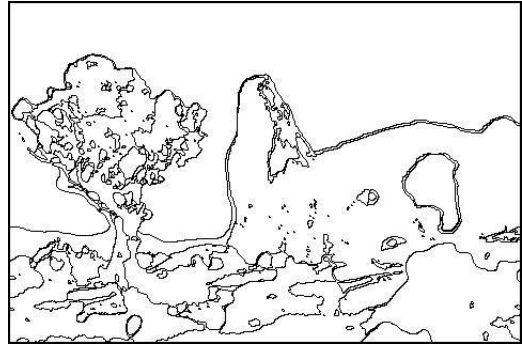
(b)



(c)



(d)



(e)

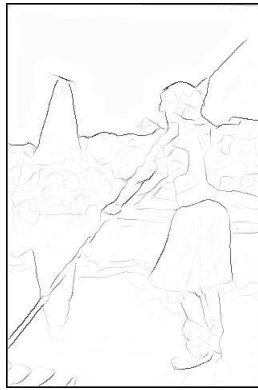
Figura 4.34. (a) Imagen original, (b) bordes obtenidos usando segmentación por *mean shift*, (c) bordes obtenidos usando algoritmo de gradiente de textura, (d) bordes obtenidos usando algoritmo basado en gradiente de color, (e) bordes obtenidos por la nueva propuesta.



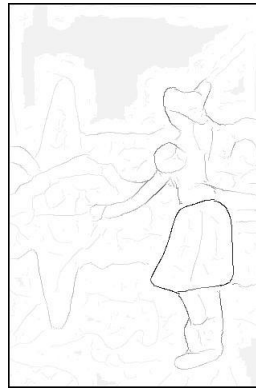
(a)



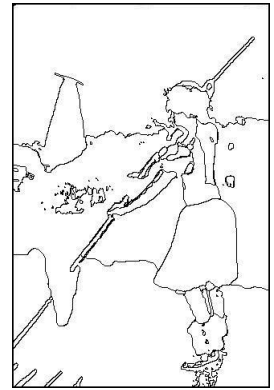
(b)



(c)



(d)



(e)

Figura 4.35. (a) Imagen original, (b) bordes obtenidos usando segmentación por *mean shift*, (c) bordes obtenidos usando algoritmo de gradiente de textura, (d) bordes obtenidos usando algoritmo basado en gradiente de color, (e) bordes obtenidos por la nueva propuesta.

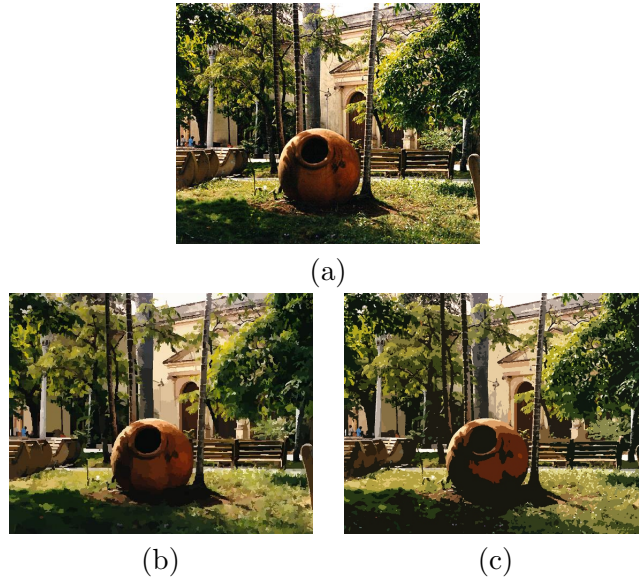


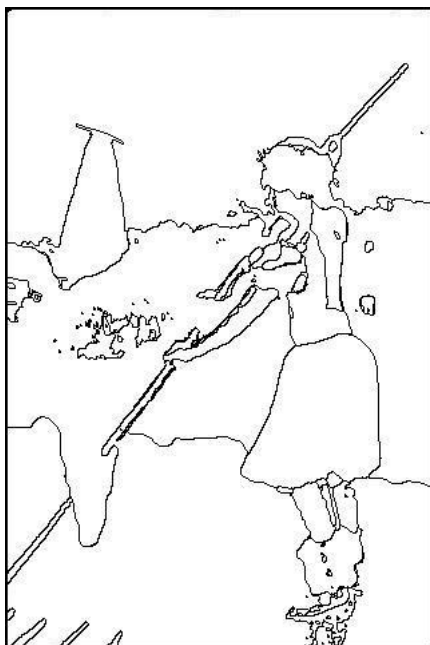
Figura 4.36. (a) Imagen original, (b) segmentación usando *mean shift*, (c) segmentación usando la nueva propuesta con representación a través de (4.3) considerando componentes conexas.

Los bordes dados por el algoritmo de *mean shift* [22] no presentan discontinuidades, sin embargo está presente una sobresegmentación. En la Figura 4.36 se muestra otro ejemplo en el que se compara la segmentación por *mean shift* con el resultado de la propuesta elaborada. Obsérvese que el uso de 11 categorías básicas de color, establecidas a través de investigaciones acerca del nombramiento del color, permite describir de forma concisa la escena en colores que se observa.

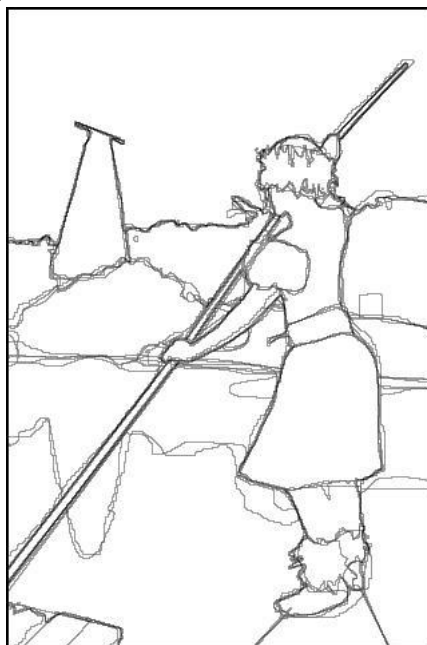
Por su parte, en los ejemplos que muestran los resultados de los algoritmos basados en gradiente de color y texturas [70], los bordes detectados presentan discontinuidades. Los resultados del método propuesto fueron comparados también con las segmentaciones manuales dadas en la base de datos de Berkeley [70]. En las Figuras 4.37 y 4.38 se muestra esta comparación. Las fronteras extraídas por la nueva propuesta destacan los elementos más sobresalientes en la imagen y están en correspondencia con el juicio humano.



(a)



(b)

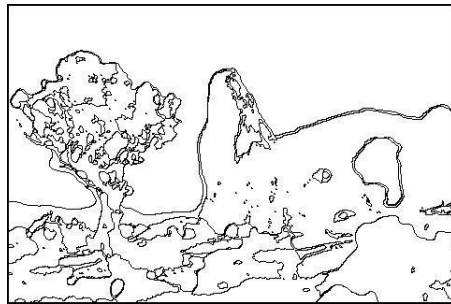


(c)

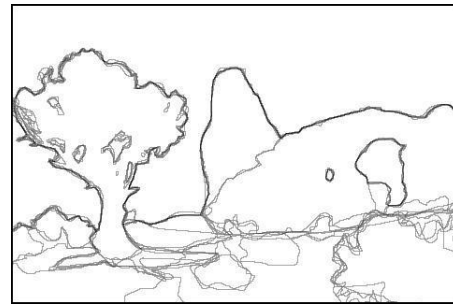
Figura 4.37. (a) Imagen original, (b) bordes detectados por la nueva propuesta, (c) segmentación manual dada en la base de datos de Berkeley.



(a)



(b)



(c)

Figura 4.38. (a) Imagen original, (b) bordes detectados por la nueva propuesta, (c) segmentación manual dada en la base de datos de Berkeley.

Capítulo 5

Otras Aplicaciones

En este capítulo se proponen dos aplicaciones de la propuesta elaborada:

1. Filtrado de ruido en imágenes en colores.
2. Uso del refinamiento para compresión.

5.1. Filtrado de ruido en imágenes en colores

Una vez calculados los atributos C y S mediante el método de segmentación propuesto, se obtienen regiones o segmentos de color en la imagen a estudiar identificadas con la etiqueta k . Las fronteras de dichas regiones serán utilizadas como barreras para realizar el filtrado de ruido en imágenes en colores, que a continuación se propone. El algoritmo consiste en realizar una difusión de la información de color relativa a los píxeles ubicados en una misma región de color, respetando las fronteras de cada una de ellas. En (5.1) se indica la ecuación a través de la cual se realiza la difusión.

$$Ch_i(r) = \frac{\sum_{s \in N_r} \delta(I(s) - k) Ch_i(s)}{\sum_{s \in N_r} \delta(I(s) - k)}, \quad r \in k. \quad (5.1)$$

En (5.1) $Ch_i(r)$ denota la información del canal i del pixel r , que es resultado de la difusión, $i \in \{L^*, u^*, v^*\}$; $Ch_i(s)$ denota la componente de color i en el sitio s . El número de iteraciones del proceso de difusión es un parámetro del algoritmo que se propone.

En las Figuras 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8 y 5.9 aparecen ejemplos de la aplicación propuesta para diferentes niveles de ruido. El ruido es de tipo Gaussiano, y fue generado en el espacio $L^*u^*v^*$; donde cada componente fue contaminada.

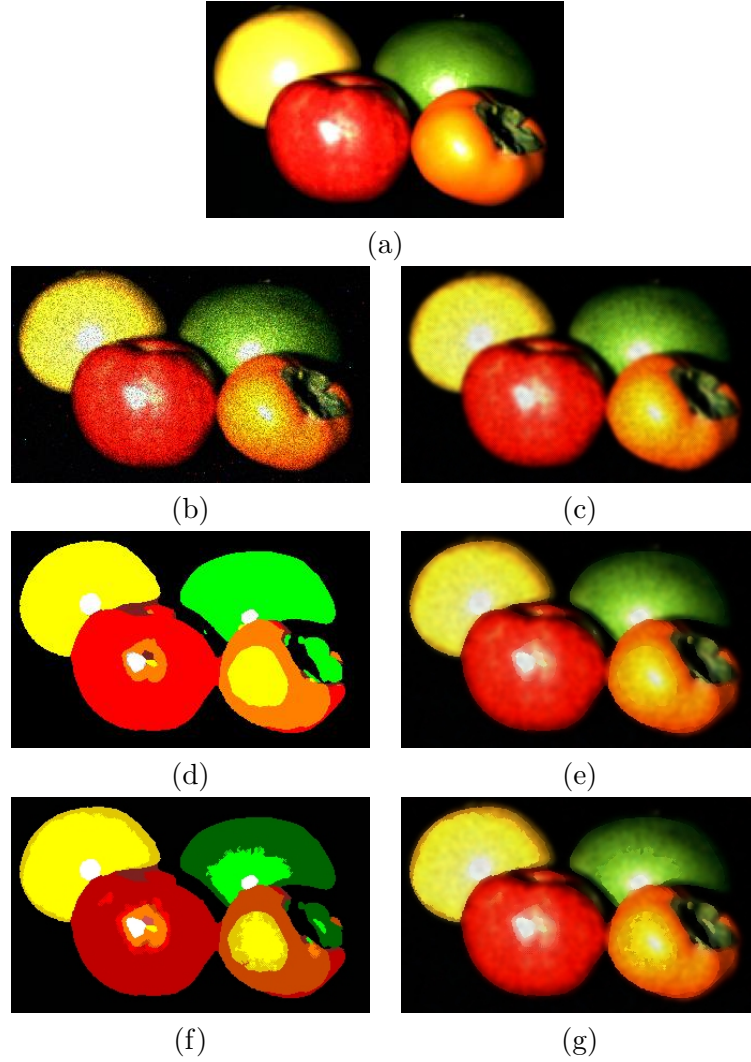


Figura 5.1. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 8$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

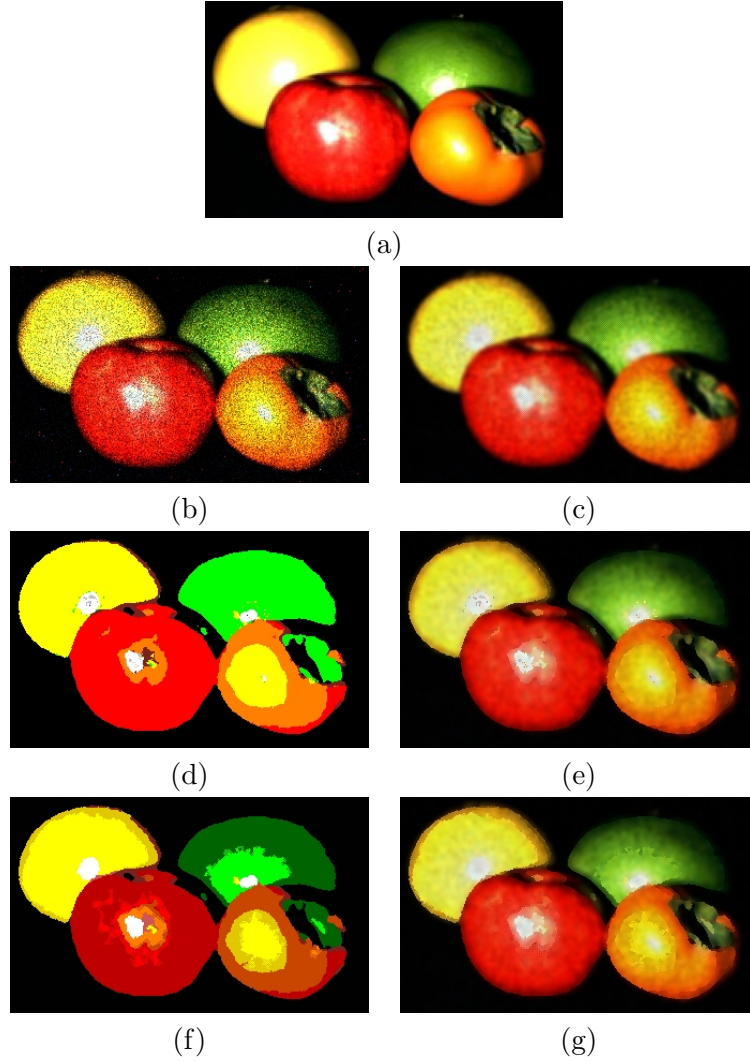


Figura 5.2. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 10$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

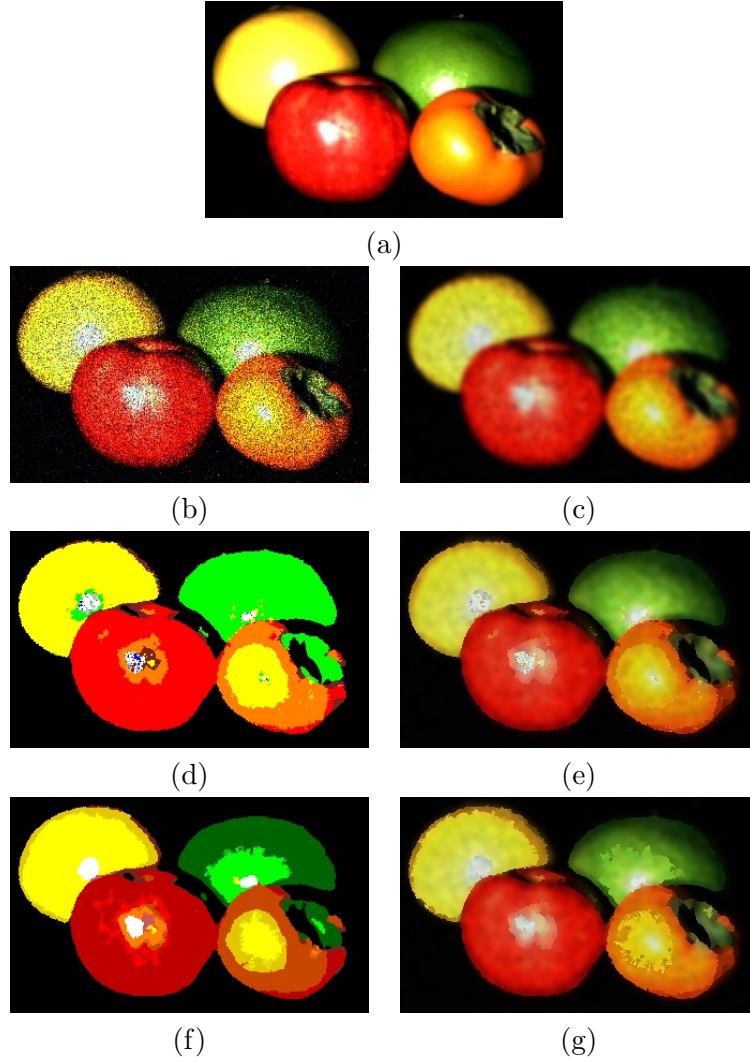


Figura 5.3. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 15$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

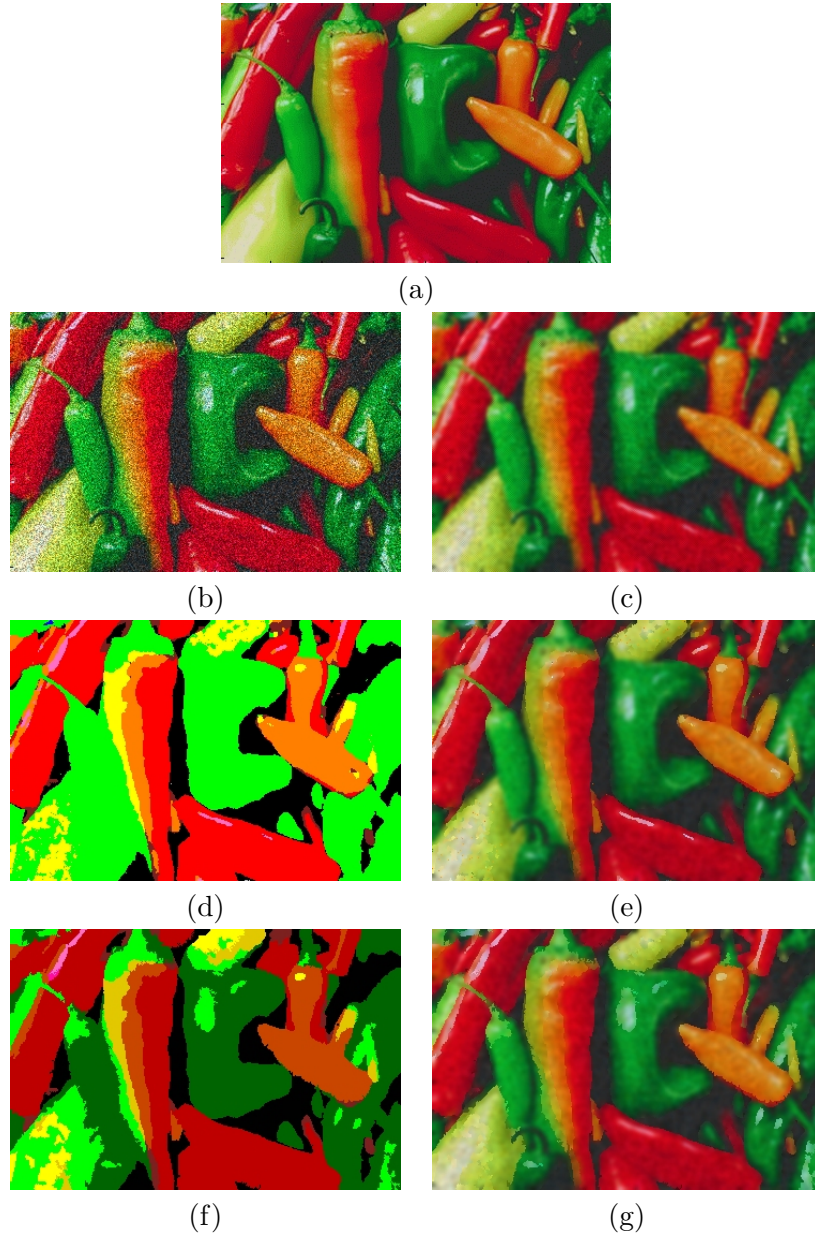


Figura 5.4. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 8$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

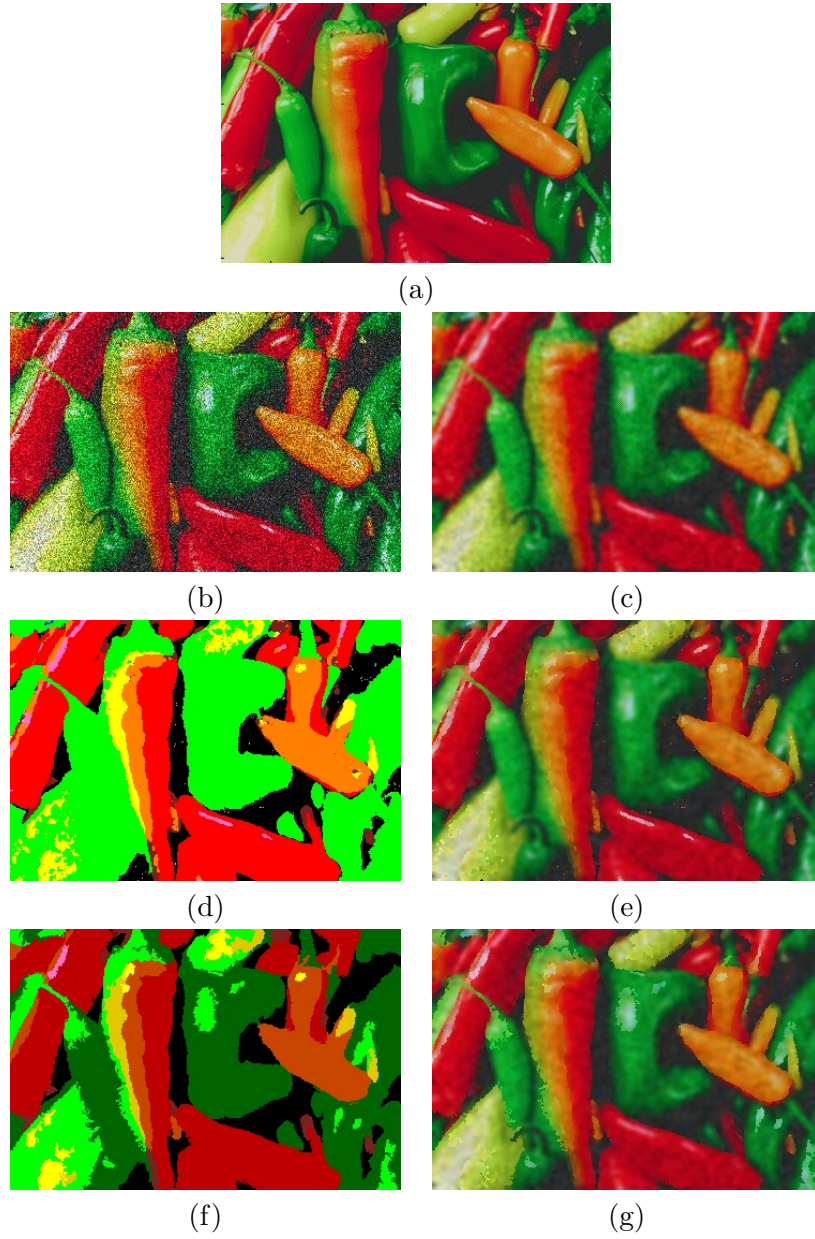


Figura 5.5. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 10$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

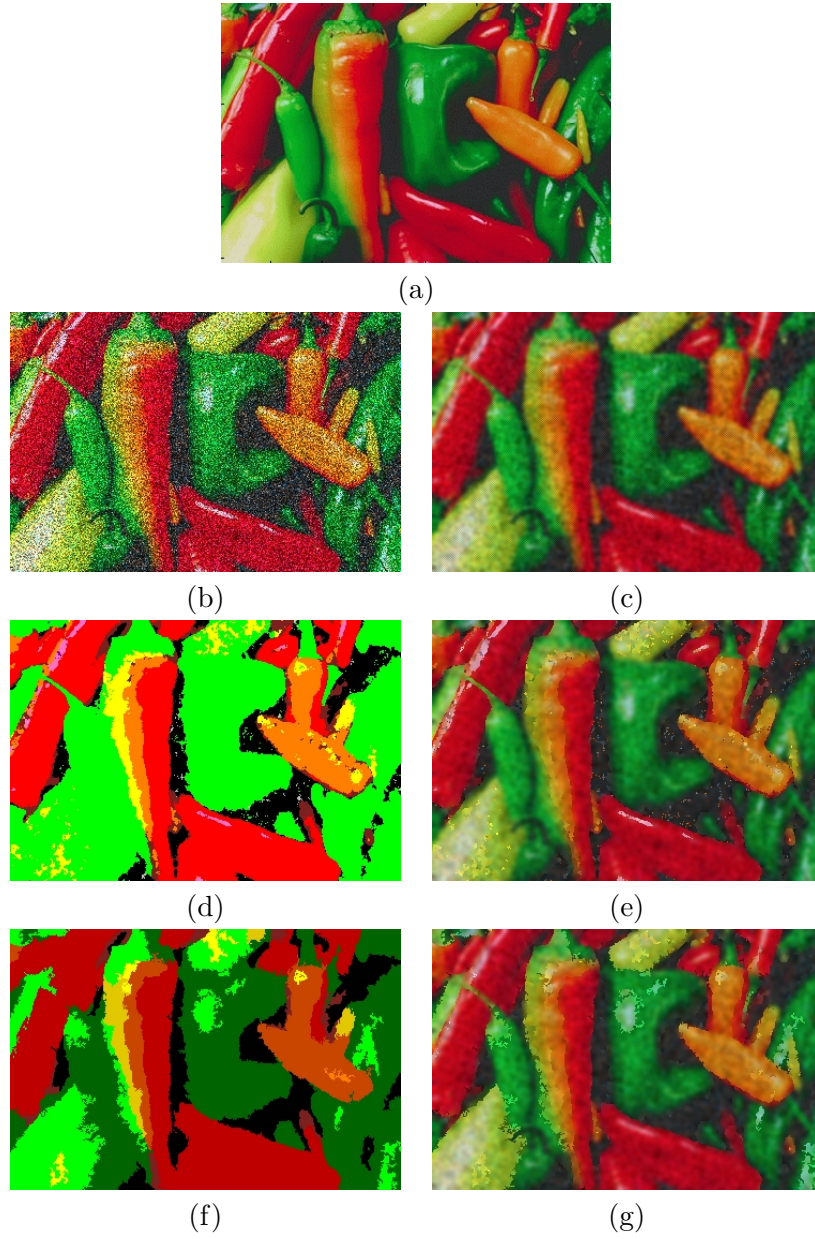


Figura 5.6. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 15$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

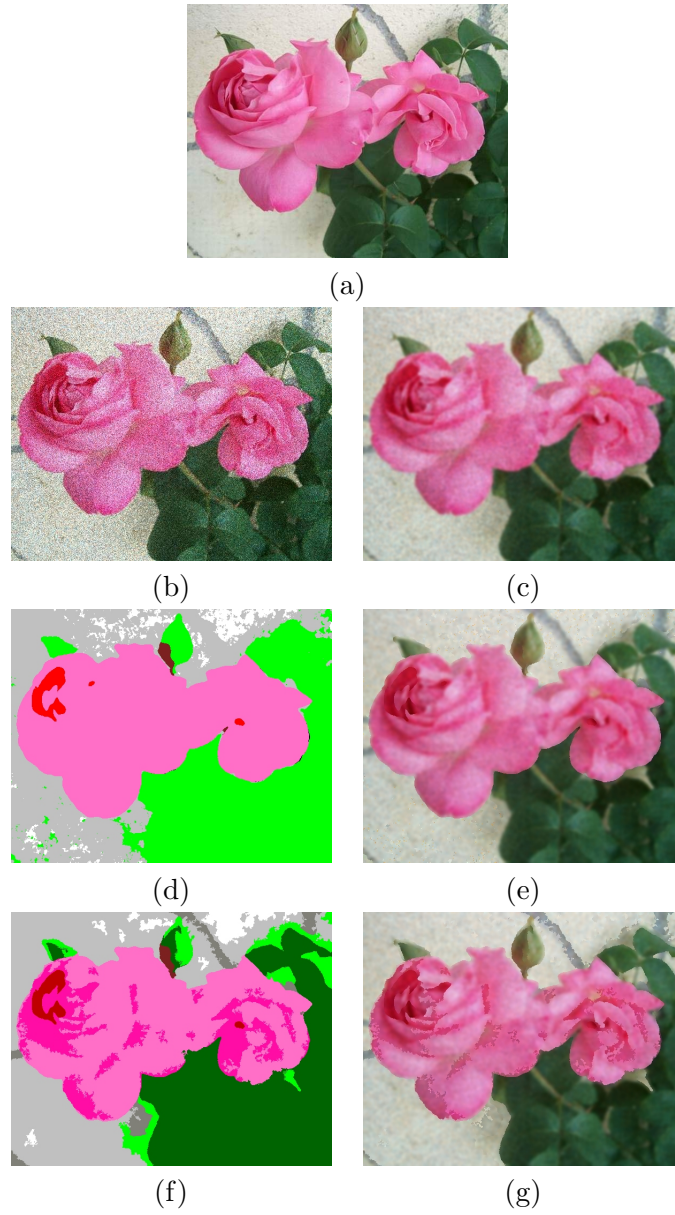


Figura 5.7. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 8$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

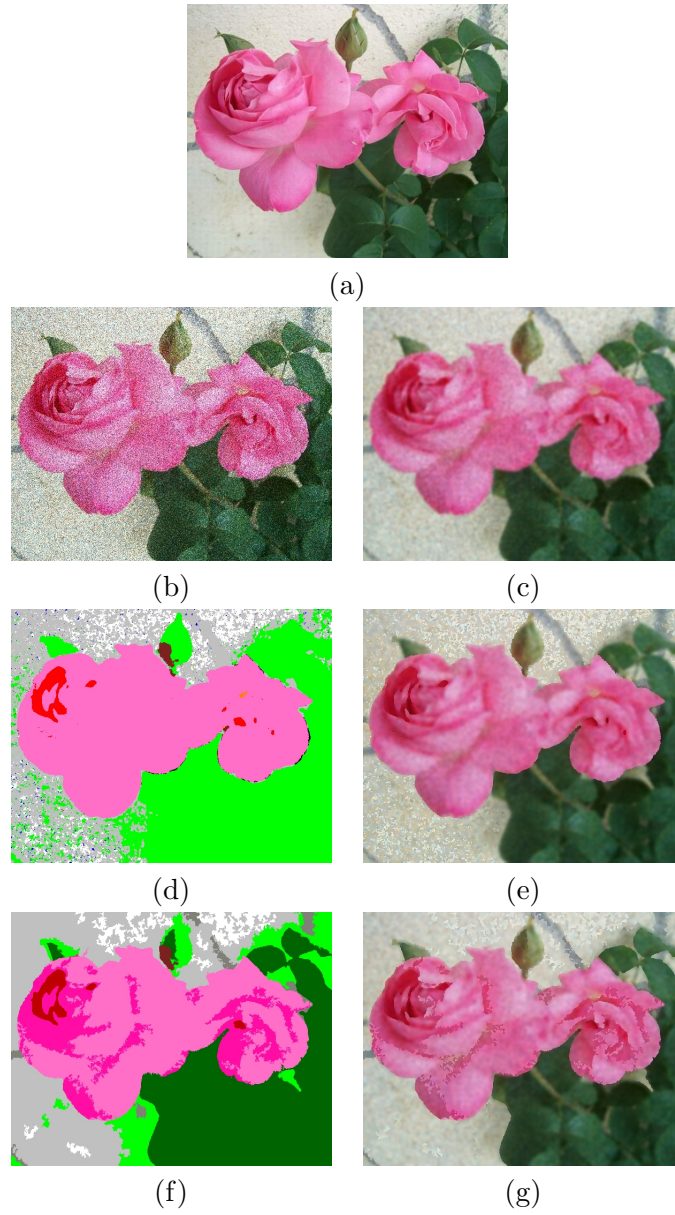


Figura 5.8. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 10$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

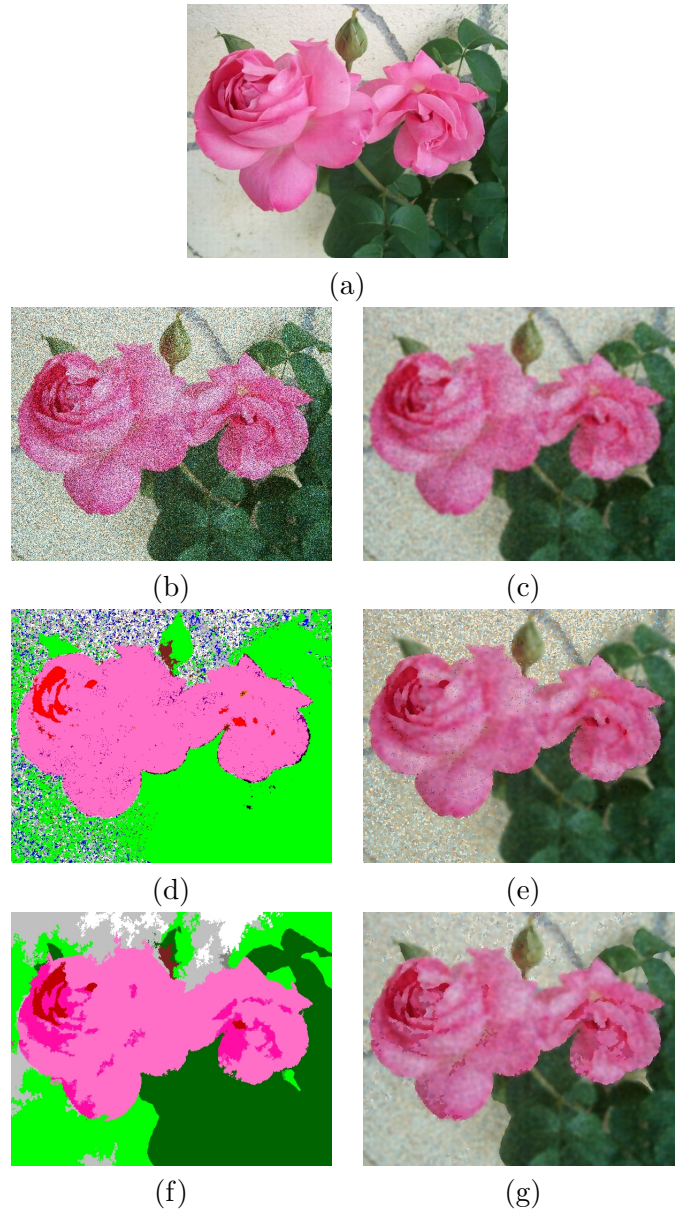


Figura 5.9. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 15$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

Obsérvese que al filtrar ruido, sin considerar los resultados de la propuesta de segmentación (Figuras 5.1(c), 5.2(c), 5.3(c), 5.4(c), 5.5(c), 5.6(c), 5.7(c), 5.8(c) y 5.9(c)) se afecta la información de contornos.

Los ejemplos mostrados parten de imágenes en las que predominan zonas homogéneas; en la Figura 5.10 se muestran otro ejemplo de la aplicación propuesta, en el que la homogeneidad en color es menor.

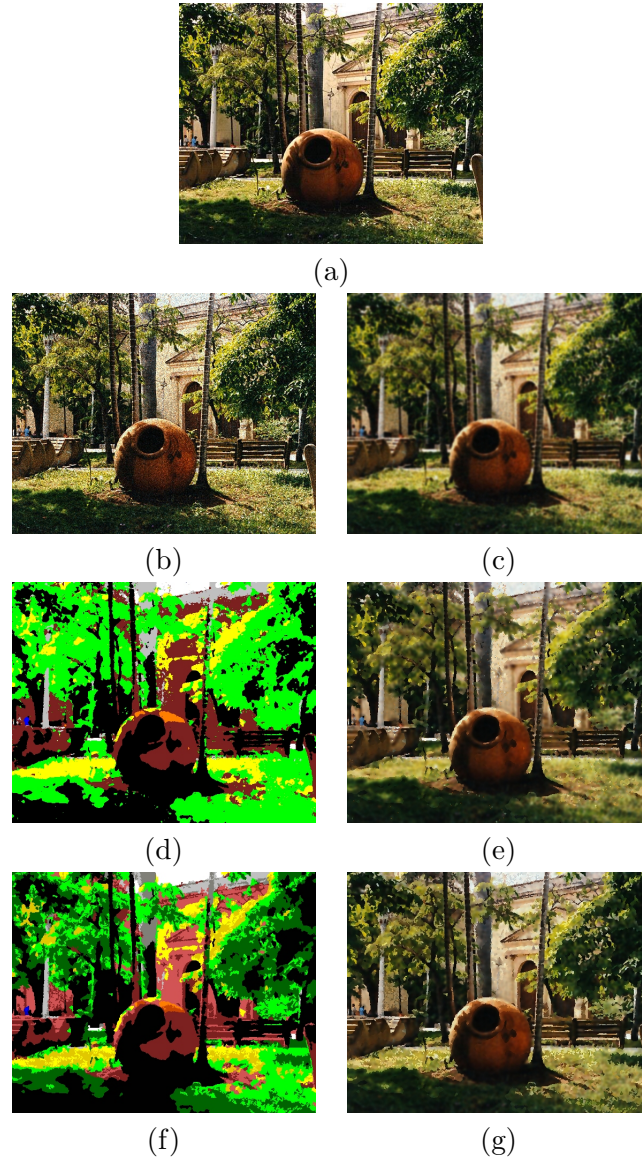


Figura 5.10. (a) Imagen original, (b) imagen contaminada con ruido Gaussiano de $\sigma = 15$, (c) imagen filtrada sin respetar bordes, (d) imagen resultado de la estimación del atributo C, (e) imagen filtrada usando información de bordes, luego de estimar el atributo C, (f) imagen resultado de la estimación del atributo S y postprocesamiento, (g) imagen filtrada usando información de bordes, luego de estimar atributo S (uso de la expresión (5.1)).

Los resultados dados en las Figuras 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 y 5.10 ilustran el comportamiento del método elaborado ante ruido. Para valores mayores de σ la distorsión del color es muy grande y la estimación del atributo C no logra encontrar las categorías de color correctas. Los valores adecuados de λ estuvieron en el intervalo 8-10; el valor usado para μ fue 0.5. En la aplicación propuesta no se consideran las interacciones entre colores.

5.2. Compresión de imágenes

La segmentación mediante la propuesta descrita da lugar a regiones de color, que perceptualmente son las más significativas en la descripción de la composición del color, en una imagen que se observa. Inspirados en lo dicho anteriormente se propone usar la salida del método propuesto en la etapa de cuantificación de algoritmos de compresión.

En la sección 3.4.1 se describió que luego de la etapa de estimación del atributo C, donde se detectan los colores básicos, definidos por Berlin y Kay [6], le sigue una etapa de refinamiento de cada categoría encontrada usando los atributos de intensidad: *claro* y *oscuro*; de esta forma cada categoría es subdividida en dos clases y es posible dar una descripción más detallada de la composición en color de la imagen. Si en lugar de dos clases para el refinamiento, se agregan más, el paso de refinamiento acercará más el resultado final de la segmentación a la imagen original. Considérense entonces un número n de subclases de refinamiento y el siguiente algoritmo para compresión basado en la propuesta dada:

Algorithm 1 Refinamiento para compresión

- 1: Estimación del atributo C usando la expresión (3.24).
 - 2: Estimación de n subclases para el atributo C.
 - 3: Representación del resultado anterior usando la expresión (4.3).
-

En el algoritmo descrito en (1) el paso 2 indica la etapa de refinamiento de cada categoría detectada en el paso 1, para dar lugar a subclases que permiten una descripción más detallada de la información de color (estimación del atributo S). En la sección (3.4.1) se explica el procedimiento de refinamiento considerando $n = 2$. La aplicación que se propone extiende la cantidad de modelos de intensidad (subclases) a un número mayor usando el mismo procedimiento.

En los ejemplos mostrados de las Figuras 5.11 y 5.12 se muestra la información de la imagen original en forma compacta, sin perder los aspectos más importantes de la original. Las imágenes están representadas a través de las zonas de color más significativas, lo que podría ser útil para la etapa de cuantificación en un algoritmo de compresión.

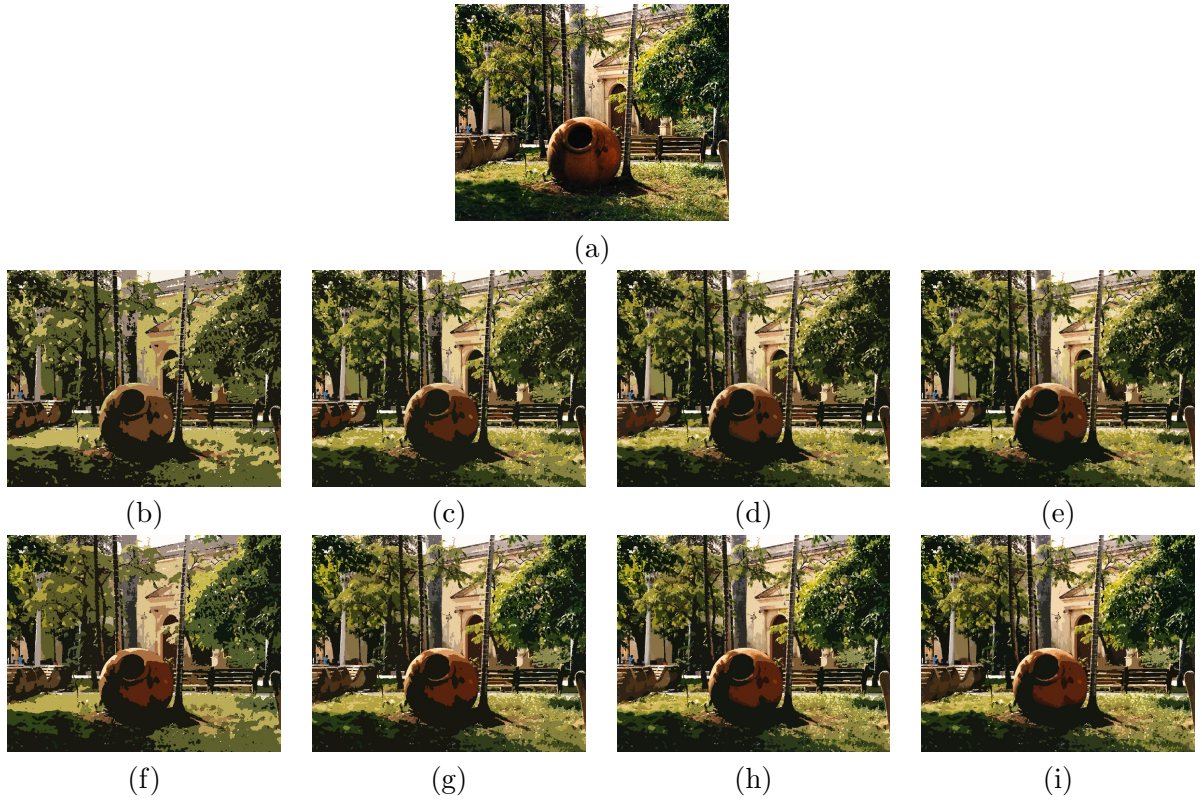


Figura 5.11. (a) Imagen original, (b) representación del atributo C. Refinamiento del atributo C a través de (c) 2 subclases, (d) 4 subclases y (e) 6 subclases usando la ecuación (4.3) en cada subclase encontrada. Refinamiento del atributo C a través de (f) 2 subclases, (g) 4 subclases y (i) 6 subclases usando la ecuación (4.3) en cada componente conexas.

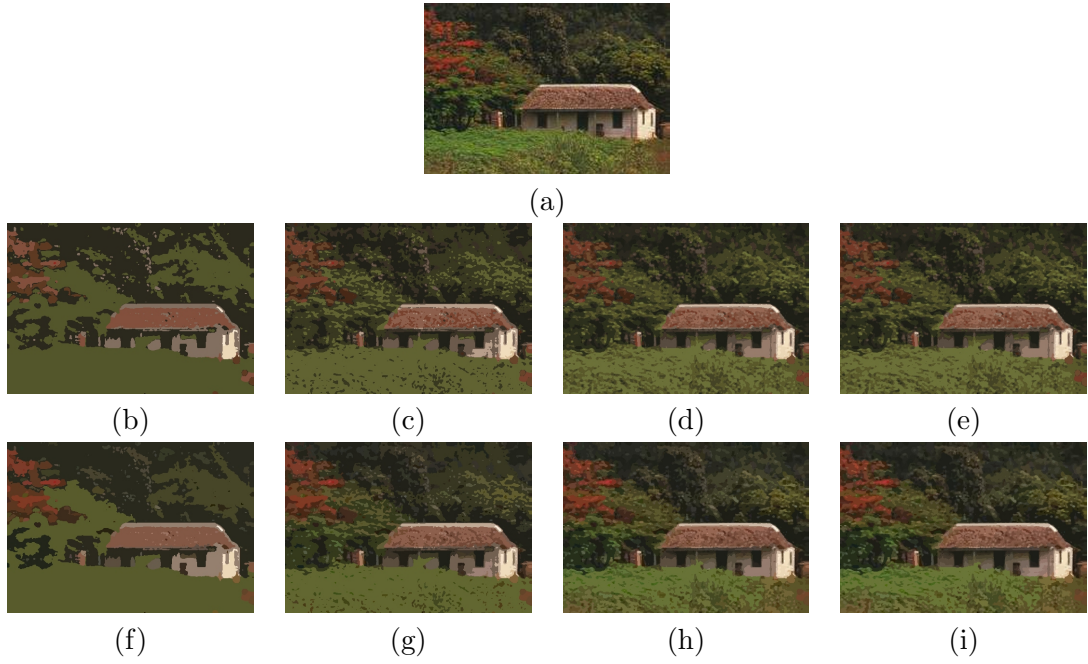


Figura 5.12. (a) Imagen original, (b) representación del atributo C. Refinamiento del atributo C a través de (c) 2 subclases, (d) 4 subclases y (e) 6 subclases usando la ecuación (4.3) en cada subclase encontrada. Refinamiento del atributo C a través de (f) 2 subclases, (g) 4 subclases y (i) 6 subclases usando la ecuación (4.3) en cada componente conexa.

Capítulo 6

Conclusiones

En este trabajo de tesis se ha elaborado un método de segmentación, en el que por primera vez se fusiona un modelo de categorización del color con técnicas Bayesianas de interacción espacial.

El modelo de categorización del color propuesto, considera las investigaciones realizadas sobre el nombramiento del color por Berlin y Kay [6] en 1969. Dicho modelo constituye una forma novedosa de interpolación, donde se combina información del espacio RGB, con el espacio $L^*u^*v^*$. Como resultado de la modelación se calcula un vector de verosimilitud para cada voxel en el espacio de color seleccionado y el conjunto de vectores da lugar a un diccionario probabilístico. Conociendo la componente de mayor valor en el vector de verosimilitudes se conoce la categoría básica de color más verosímil en cada voxel, lo que permite particionar el espacio en las 11 categorías de color definidas en [6] y saber la distribución y forma aproximada de las regiones asociadas a cada clase de color. Aunque el modelo se limita a 11 categorías de color, su extensión a otro vocabulario es inmediata, para lo cual sólo se necesita cambiar la etapa de entrenamiento.

El diccionario probabilístico obtenido es usado en combinación con técnicas bayesianas de interacción espacial, haciéndose una generalización de la propuesta dada en [30], que permite considerar interacciones entre las categorías básicas de color, pero puede usarse para otros propósitos. En esta propuesta de tesis las interacciones entre colores consideradas se basan en la relación perceptual entre categorías, establecida experimentalmente en [26].

La segmentación de imágenes en colores, que se propone, permite determinar un vector de probabilidades por cada pixel, en el que cada componente indica su grado de pertenencia a cada una de las 11 clases de color (paso de estimación del atributo C). Al calcular la moda de la distribución se tiene entonces la clase a la que pertenece ese pixel. Un aspecto importante de

esta etapa es el manejo de la incertidumbre en zonas oscuras de color, dando lugar a realzar el valor de las componentes asociadas al color cuando sea posible.

Con la propuesta de segmentación se calculan además los atributos de intensidad claro y oscuro de cada región de color encontrada (paso de estimación del atributo S). Esta etapa de refinamiento puede extenderse a más de dos subclases, permitiendo así la extracción de un número mayor de detalles en la imagen. El resultado de la propuesta de segmentación extrae las regiones de color más significativas, aspecto éste que se propone usar en la etapa de cuantificación para compresión de imágenes en colores.

Finalmente el proceso descrito da lugar a la partición de la imagen en regiones de color, de forma similar al observador humano.

Con el método elaborado se obtiene un mapa de bordes que se corresponde con la información de contorno más sobresaliente, acorde al juicio humano. Este resultado es usado como guía en la realización de un filtrado preservando bordes; aplicación esta que se propone.

Como aspecto muy importante es el hecho de procesar la información de color como un todo al segmentar la imagen y la selección de un número de clases no arbitrario, sino basado en investigaciones relevantes del proceso humano de nombramiento del color.

Para hablar del trabajo futuro es necesario señalar las debilidades del trabajo realizado:

1. No se considera la constancia de color: aunque en el método elaborado se simula la incertidumbre presente en zonas oscuras, esto es sólo un primer paso en la modelación de la constancia del color.
2. Las interacciones entre colores incluidas se limitan al vocabulario seleccionado, y se basan en experimentos hechos con ese vocabulario.
3. Aunque los espacios $L^*a^*b^*$ y $L^*u^*v^*$ demuestran ser adecuados para el procesamiento del color, no se examinó el comportamiento del método en espacios como el HSI, reportado como un espacio, que simula correctamente la descripción humana del color empleando los atributos de tono, saturación e intensidad.

Apéndice A

Sobre splines

Un spline es una función definida por una familia de polinomios “sociables”, donde el término sociable indica la vinculación que existe entre ellos. La palabra *spline*, del idioma inglés, viene de un instrumento usado por ingenieros navales para dibujar curvas suaves, forzadas a pasar por un conjunto de puntos prefijados [72].

En general cualquier polinomio $P(x)$ de grado menor o igual que n se puede expresar de la forma:

$$P(x) = \alpha_1 P_1(x) + \alpha_2 P_2(x) + \alpha_3 P_3(x) + \dots + \alpha_n P_n(x). \quad (\text{A.1})$$

donde $\{P_1(x), P_2(x), \dots, P_n(x)\}$ es un conjunto de polinomios independientes, es decir ninguno de ellos se puede expresar como combinación lineal de los otros. La familia de polinomios es usada para resolver problemas de aproximación, específicamente el caso de interpolación. Esto se debe entre otras razones a la siguiente propiedad [72]: Dados $n + 1$ puntos distintos (x_i, y_i) existe un único polinomio $p(x)$ de grado menor o igual que n tal que $P(x_i) = y_i$ $i = 1, 2, \dots, n$. En la Figura A.1 aparecen algunos ejemplos de aproximaciones de curvas dado un número determinado de puntos. A pesar de la propiedad antes señalada, es frecuente que para ciertas distribuciones de datos se tenga un polinomio cuyo comportamiento no es el deseado debido al número de oscilaciones que tiene (ver Figura A.1(d)). Para eliminar este problema, en lugar de usar un solo polinomio se usa una familia de polinomios “sociables”. Los splines cuadráticos son un ejemplo de esta familia. A continuación se explica como se construyen los mismos.

Sea:

$$S_0 = \{s : [a, b] \rightarrow \mathbb{R} \mid s(x) = P_1(x), x \in I_1, s(x) = P_2(x), x \in I_2\}, \quad (\text{A.2})$$

$$I_1 = [a, \xi], I_2 = [\xi, b], P_1(x) \text{ y } P_2(x) \text{ son polinomios de grado 2.} \quad (\text{A.3})$$

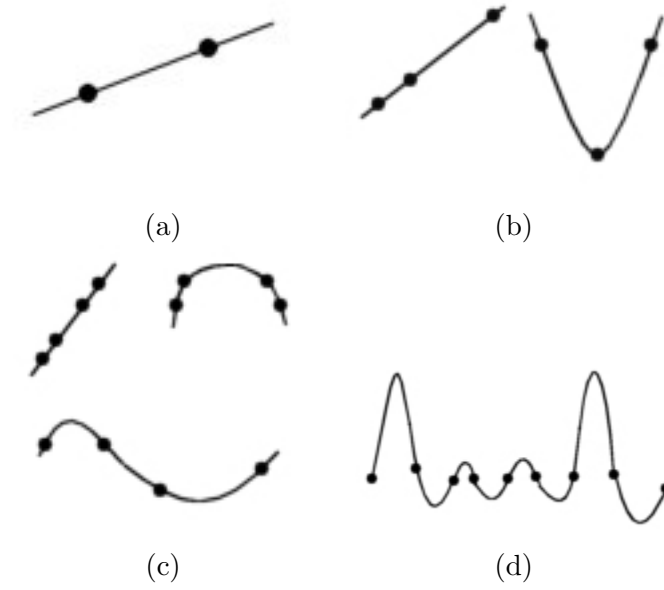


Figura A.1. Ejemplos de uso de splines dados a) dos puntos, b) tres puntos, c) 4 puntos, d) n puntos

S_0 es lineal. Considérese además el subconjunto de S_0 definido por las funciones continuas de este espacio:

$$S_1 = \{s(x) \in S_0 \mid P_1(\xi) = P_2(\xi)\}. \quad (\text{A.4})$$

Sea además el subconjunto de funciones continuamente diferenciable definido como:

$$S_2 = \{s(x) \in S_1 \mid P_1'(\xi) = P_2'(\xi)\}. \quad (\text{A.5})$$

S_1 es un subespacio lineal de S_0 y S_2 es un subespacio lineal de S_1 . La función $s(x)$ que cumple las condiciones descritas se conoce como spline cuadrático. En la propuesta de tesis se usa el spline cuadrático definido por la siguiente expresión:

$$\beta(x) = \begin{cases} \frac{1}{2}(-2x^2 + 1,5), & |x| \in \left[0, \frac{1}{2}\right] \\ \frac{1}{2}(x^2 - 3|x| + 2,25), & |x| \in \left[\frac{1}{2}, 1,5\right] \\ 0, & |x| \geq 1,5 \end{cases} \quad (\text{A.6})$$

Apéndice B

Experimentos

A continuación se muestran otros ejemplos de la segmentación en imágenes en colores, usando la propuesta elaborada. En la Figura B.1 se especifica la leyenda empleada al codificar los atributos C y S.


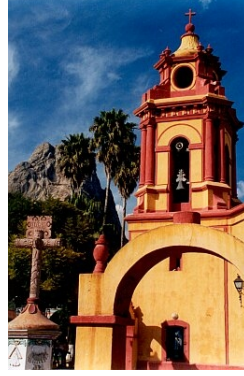
	rojo oscuro
	rojo (rojo claro)
	verde oscuro
	verde (verde claro)
	azul oscuro
	azul (azul claro)
	amarillo oscuro
	amarillo (amarillo claro)
	púrpura oscuro
	púrpura (púrpura claro)
	anaranjado oscuro
	anaranjado (anaranjado claro)
	gris oscuro
	gris (gris claro)
	negro
	blanco
	rosa oscuro
	rosa (rosa claro)
	(café) café oscuro
	café claro

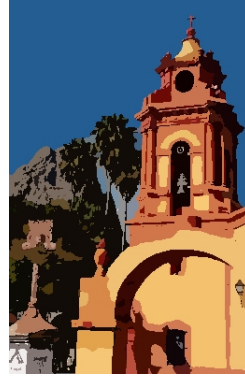
Figura B.1. Leyenda empleada para la codificación de los atributos C y S durante la segmentación.



(a)



(b)



(c)



(d)

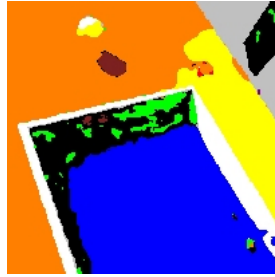


(e)

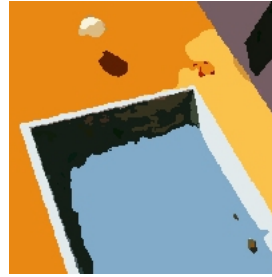
Figura B.2. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



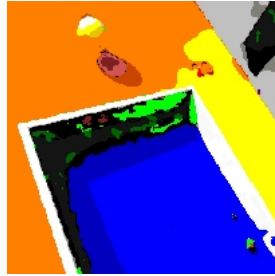
(a)



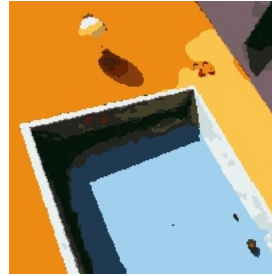
(b)



(c)



(d)



(e)

Figura B.3. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



(a)



(b)



(c)



(d)



(e)

Figura B.4. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



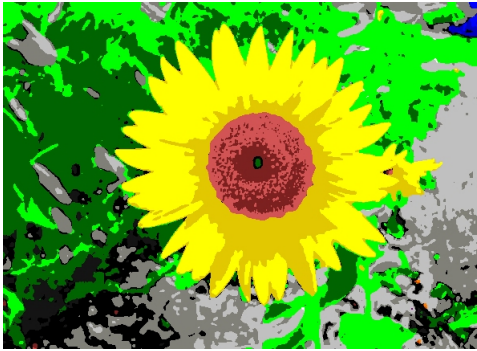
(a)



(b)



(c)



(d)

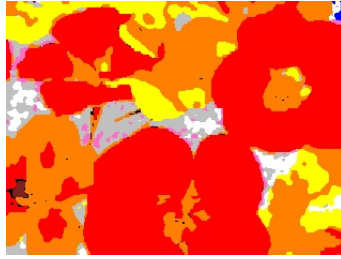


(e)

Figura B.5. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



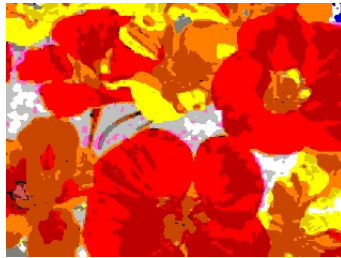
(a)



(b)



(c)



(d)

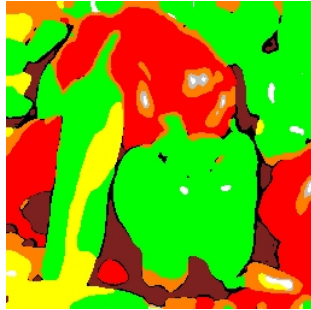


(e)

Figura B.6. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



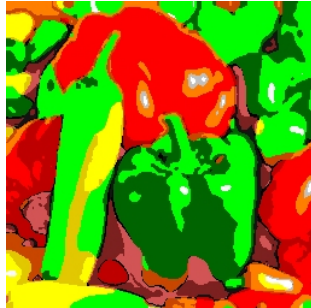
(a)



(b)



(c)

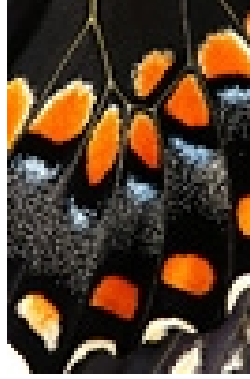


(d)



(e)

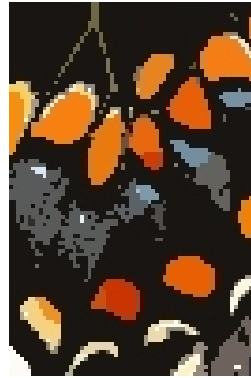
Figura B.7. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



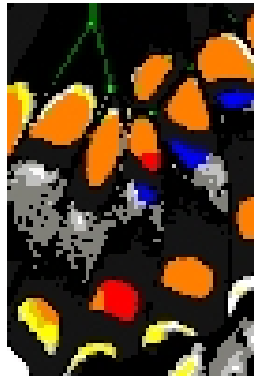
(a)



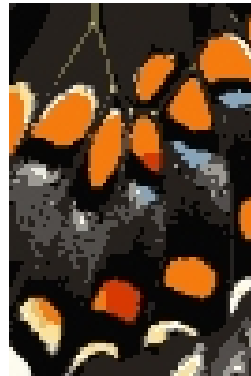
(b)



(c)



(d)

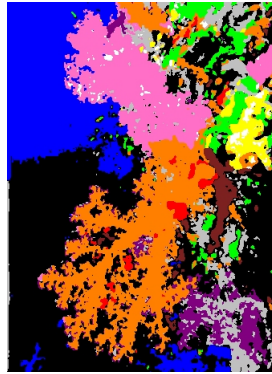


(e)

Figura B.8. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



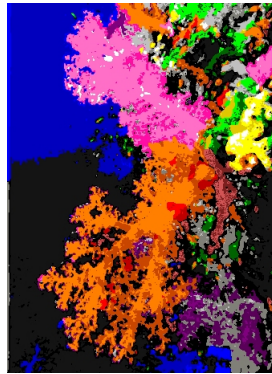
(a)



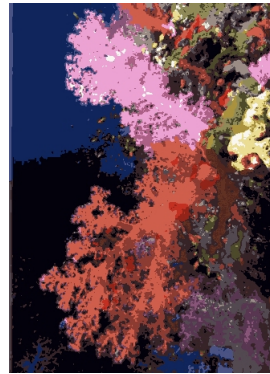
(b)



(c)



(d)

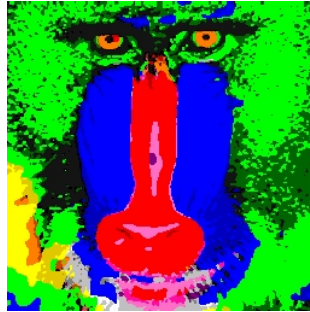


(e)

Figura B.9. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



(a)



(d)

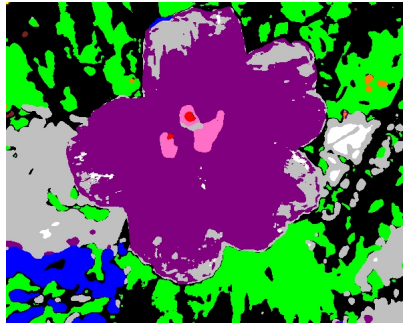


(e)

Figura B.10. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



(a)



(b)



(c)



(d)

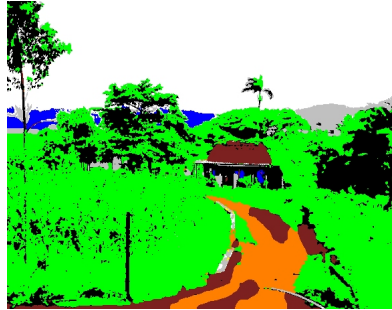


(e)

Figura B.11. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



(a)



(b)



(c)



(d)



(e)

Figura B.12. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).



(a)



(b)



(c)



(d)



(e)

Figura B.13. (a) Imagen original, (b) estimación del atributo C, (c) representación del atributo C usando expresión (4.3), (d) estimación del atributo S dado C, (e) representación del atributo S usando expresión (4.3).

Bibliografía

- [1] H. V. Helmholtz, *Physiological optics*, 3rd ed.
- [2] E. Hering, “Principles of a new theory of color sense,” *Color Vision*, 1961.
- [3] R. Kenneth, J. Sloan, and R. Bajcsy, “A computational structure for color perception,” in *The ACM digital library, proceedings of the 1975 annual conference*, January 1975.
- [4] R. L. D. Valois, I. Abramov, and G. H. Jacobs, “Analysis of response patterns of lgn cells,” *Journal of the Optical Society of America*, vol. 56, no. 7, pp. 966–977, October 1966.
- [5] A. Mojsilovic, “A method for color naming and description of color composition in images,” *IEEE Transactions on Image Processing*, vol. 14, no. 5, pp. 690–699, May 2005.
- [6] B. Berlin and P. Kay, *Basic Color Terms: Their universality and Evolution*. Berkeley: University of California, 1969.
- [7] C. McDaniel, “Hue perception and hue naming,” Ph.D. dissertation, Harvard college, 1972.
- [8] H. Cheng, X. Jiang, Y. Sun, and J. Wang, “Color image segmentation: advances and prospects,” *Pattern Recognition*, vol. 34, pp. 2259–2281, June 2001.
- [9] S. P. et. al., “A review on image segmentation techniques,” *Pattern Recognition*, vol. 29, pp. 1277–1294, 1993.
- [10] T. L. Huntsberger, C. Jacobs, and R. L. Cannon, “Iterative fuzzy image segmentation,” *Pattern Recognition*, vol. 18, no. 2, pp. 131–138, 1985.
- [11] J. M. Pardo, A. F. Solé, and D. Cabello, “Active segmentation of biomedical images: edge saliency and region dissimilarity,” *Pattern Recognition and Image Analysis, Proceedings of the VIII national symposium in Shape recognition and Image analysis*, vol. 1, pp. 513–520, 1999.

- [12] H. D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Trans. Image Processing*, vol. 9, no. 12, pp. 2071–2082, 2000.
- [13] J.-F. Yang, S.-S. Hao, and P.-C. Chung, "Color image segmentation using fuzzy c-means and eigenspace projections," *Signal Processing*, vol. 82, pp. 461–472, 2002.
- [14] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden markov random field and the expectation maximization algorithm," *IEEE Trans. Medical Imaging*, vol. 20, no. 1, January 2001.
- [15] T. Uchiyama and M. A. Arbib, "Color image segmentation using competitive learning," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 16, no. 12, pp. 1197–1206, 1994.
- [16] F. Ferri and E. Vidal, "Colour image segmentation and labeling through multiedit-condensing," *Pattern Recognition Letters*, vol. 13, pp. 561–568, 1992.
- [17] C. Huang, T. Cheng, and C.-C. Chen, "Color images segmentation using scale filter and Markov random field," *Pattern Recognition*, vol. 25, no. 10, pp. 1217–1229, 1992.
- [18] P. Schmid, "Segmentation of digitized dermatoscopic images by two-dimensional color clustering," *IEEE Trans. Medical imaging*, vol. 18, No. 2, 164–171, February 1999.
- [19] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham, "Classification and localisation of diabetic-related eye disease," *ECCV 2002, LNCS 2353*, pp. 502–516, 2002.
- [20] J. C. Rajapakse and F. Kruggel, "Segmentation of MR images with intensity inhomogeneities," *Image and Vision computing*, vol. 16, pp. 165–180, 1998.
- [21] S. Lefevre, L. Mercier, V. Tiberghien, and N. Vincent, "Multiresolution color image segmentation applied to background extraction in outdoor images," *RFAI Publication, ECCV, Poitiers France*, pp. 363–367, april 2002.
- [22] D. Comaniciu and P. Meer, "Mean shift analysis and applications," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [23] R. Nock and F. Nielsen, "Statistical region merging," *IEEE on PAMI*, vol. 26, pp. 1452–1458, 2004.
- [24] R. C. González and R. E. Woods, *Tratamiento Digital de Imágenes*, 2nd ed., J. Goodman, Ed. Massachusetts, E.U.A: Addison-Wesley, 1996.

- [25] C. Huang, "Pattern image segmentation using modified Hopfield model," *Pattern Recognition Lett.*, no. 13, pp. 345–353, 1999.
- [26] R. M. Boynton and C. X. Olson, "Locating basic colors in the OSA space," *Color Res Appl*, vol. 12, pp. 94–105, 1987.
- [27] J. E. Cairo, "The neurophysiological basis of basic color terms," Ph.D. dissertation, State University of New York at Binghamton, 1977.
- [28] P. Kay and C. McDaniel, "The linguistic significance of meanings of basic color terms," *Color Res Appl*, vol. 54, pp. 610–646, 1978.
- [29] J. Sturges and T. Whitfield, "Locating basic colors in the Munsell space," *Color Res Appl*, vol. 20, pp. 364–376, 1995.
- [30] M. Rivera, O. Ocegueda, and J. L. Marroquín, "Entropy controlled gauss-markov random measure field models for early vision," *VLSM, LNCS 3752*, pp. 137–148, 2005.
- [31] G. Wyszecki and W. Stiles, *Color Science: Concepts and methods, quantitative data and formulae*, 2nd ed., Wiley, Ed. Wiley, 1982.
- [32] N. Vandnbroucke, L. Macaire, and J.-G. Postaire, "Color image segmentation by pixel classification in adapted hybrid color space. application to soccer image analysis." *Computer Vision and Image Understanding*, vol. 10, no. 2, pp. 190–216, 2003.
- [33] G. H. Hutson, *Teoría de la television en color*, McGrawHill, Ed., 1985.
- [34] S. Tominaga, "Expansion of color images using three perceptual attributes," *Pattern Recognition Lett.*, vol. 6, pp. 77–85, 1987.
- [35] T. Alarcón, "Segmentación de imágenes en colores. su empleo en el estudio de la angiogénesis," Tesis de Maestría, Instituto Superior Politécnico José Antonio Echeverría, La Habana, Cuba, Diciembre 1998.
- [36] W. D. Wright, "A re-determination of the trichromatic coefficients of spectral colours," *Transactions of the Optical society*, vol. 30, 141-164, 1928.
- [37] J. Guild, "The colorimetric proprieties of the spectrum," *Philosophical transactions of the Royal Society of London*, vol. A230, 149-187, 1931.
- [38] B. LindBloom. (2002) Useful color equations. [Online]. Available: <http://www.brucelindbloom.com/>

- [39] S. S. Stevens, *Psychophysics: Introduction to its perceptual, neural and social prospects*. Wiley, 1975.
- [40] R. Ohlander, K. Price, and D. Reddy, "Picture segmentation using a recursive region splitting method," *Comput. Graphics Image Process.*, vol. 8, pp. 935–952, 1978.
- [41] D. C. Tseng and C. Chang, "Color segmentation using perceptual attributes," in *IEEE Internacional Conference on Pattern Recognition*, 1992, pp. 228–231.
- [42] Y. Otha, T. Kanade, and T. Sakai, "Color information for region segmentation," *Comput. Graphics Image Processing*, vol. 13, pp. 22–241, 1980.
- [43] K. Fukunaga, *Introduction to statistical pattern recognition*. Academic Press, 1990.
- [44] Y. W. Lim and S. U. Lee, "On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques," *Pattern Recognition*, vol. 23, no. 9, pp. 935–952, 1990.
- [45] J. Fan, D. K. Y. Yau, A. K. Elmagarmid, and W. G. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," *IEEE Transaction on Image Processing*, vol. 10, pp. 1454–1466, October 2001.
- [46] F. Perez and C. Koch, "Toward color image segmentation in analog VLSI: algorithm and hardware," *Int. J. Comput. Vision*, vol. 12, no. 1, pp. 17–42, 1994.
- [47] J. Canny, "A computational approach to edge detection," *IEEE Transac. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, November 1986.
- [48] D. Sinclair, "Voronoi seeded colour image segmentation," AT&T, Labs Cambridge, Tech. Rep., 1999.
- [49] J. J. Hopfield and D. W. Tank, "Neural computation of decisions in optimization problems," *Biolog. Cybernetics*, vol. 52, pp. 141–152, 1985.
- [50] G. T. Herman and B. M. Carvalho, "Multiseeded segmentation using fuzzy connectedness," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 5, pp. 460–474, 2001.
- [51] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on PAMI*, vol. 23, no. 237, pp. 1222–1239, November 2001.

- [52] S. Geman and D. Geman, “Stochastic relaxation, Gibbs distributions and bayesian restoration of images,” *IEEE-PAMI*, no. 6, pp. 721–741, 1984.
- [53] S. Z. Li., *Markov random field modeling in image analysis*. Springer-Verlag Tokyo, 2001.
- [54] J. L. Marroquín, F. Velazco, M. Rivera, and M. Nakamura, “Gauss-Markov measure field models for low-level vision,” *IEEE-PAMI*, no. 23, pp. 337–348, 2001.
- [55] J. L. Marroquín, E. Arce, and S. Botello, “Hidden Markov measure field models for image segmentation,” *IEEE-PAMI*, no. 25, pp. 1380–1387, 2003.
- [56] (2004) Ishihara test for color blindness. [Online]. Available: <http://www.toledobend.com/colorblind/Ishihara.html>
- [57] B. R. Wooten, “The effects of simultaneous and succesive chromatic constraint on spectral hue,” Ph.D. dissertation, Brown University, 1970.
- [58] J. M. Lammens, “A computacional model of color perception and color naming,” Ph.D. dissertation, Buffalo.University of New York, 1994.
- [59] R.Ñ. Shepard, “Toward a universal law of generalization for psychological science,” *Science*, no. 237, pp. 1317–1323, 1987.
- [60] K. Kelly and D. Judd, “The ISCC-NBS color names dictionary and the universal color language (the iscc-nbs method of designeting colors and a dictionary of color names),” vol. NBS circular 553, November 1955.
- [61] R. Benavente, M. Vanrell, and R. Baldrich, “Estimation of fuzzy sets for computacional colour categorization,” *Color Res. Appl.*, vol. 29, pp. 5342–353, 2004.
- [62] M. Seaborn, L. Hepplewhite, and J. Stonham, “Fuzzy colour category map for the measurement of colour similarity and dissimilarity,” *Pattern recognition*, no. 38, pp. 165–177, 2005.
- [63] A. Tversky, “Features of similarity,” *Psychol. Rev.*, no. 84, pp. 327–352, 197.
- [64] S. Mattew, H. Lee, and S. John, “Pisaro: perceptual colour and textura queries using stackable mosaic,” in *The Internacional Conference on Multimedia Computing and System*, 1999.

- [65] A. Mojsilovic, J. Gomes, and B. Rogowitz, “Semantic-friendly indexing and quering of images base don the extraction of the objective semantic cues,” *Internacional journal of computer vision*, vol. 56, pp. 79–107, 2004.
- [66] A. Jain and R. C. Dubes, *Algorithm for clustering data*. NJ: Prentice hall, 1988.
- [67] E. R. A. Santana, “Modelos de campos ocultos markovianos con aplicación a la estimación de la disparidad estereoscópica,” Ph.D. dissertation, Centro de Investigación en Matemáticas México, 2004.
- [68] R. O. Duda and P. E. Hart, *Pattern classification and scene analysis*. E.U.A: John Wiley & Sons, 1973.
- [69] J. D. Mollon, J. Pokorny, and K. Knoblauch, *Normal and defective colour vision*. New York, E.U.A: Oxford University, 2003.
- [70] (2003) The berkeley segmentation dataset and benchmark. [Online]. Available: <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>
- [71] D. R. Martin, C. C. Fowlkes, and J. Malik, “Learning to detect natural image boundarries using local brightness, color, and texture cues,” *IEEE-PAMI*, vol. 26, no. 1, January 2004.
- [72] P. B. Sánchez, V. H. Mederos, and C. D. Sánchez, *El ABC de los splines*. México: Aportaciones matemáticas, 1995.