

# EXTRACCIÓN DE CARACTERÍSTICAS EN UN ESQUEMA DE RECUPERACIÓN DE IMÁGENES

J. M. FUERTES, M. LUCENA, N. PÉREZ DE LA BLANCA\*, J. FDEZ-VALDIVIA\*

Departamento de Informática. Escuela Politécnica Superior de Jaén. Universidad de Jaén

\*Departamento de Ciencias de la Computación e I.A.ETSI Informática. Universidad de Granada

Para comunicación con los autores: jmf@apolo.ujaen.es

## Resumen

En este trabajo se presentan dos módulos fundamentales dentro del esquema diseñado de recuperación de imágenes de una base de datos a partir del color y la forma de los objetos presentes en las escenas. De una parte, se diseña un nuevo método para la detección de fronteras en imágenes en color. Se muestra una aproximación al espacio perceptual (H,S,I) (un espacio de Escala de Cromaticidad Uniforme (UCS)) sobre el que se describen tanto sus propiedades como una métrica para trabajar en el mismo. De otra, se lleva a cabo un proceso de simplificación de la información tendente a formar una estructura gráfica en la que los píxeles de la imagen se agrupan entorno a regiones de color similar. Para ello se diseña un mecanismo de crecimiento de regiones en base a la información de gradiente proporcionada por el módulo detector de fronteras, así como un módulo de división de las mismas.

## 1.- Introducción

El estudio de técnicas de acceso a grandes base de datos de documentos gráficos e imágenes a través de imágenes clave es uno de los temas que mayor desarrollo ha tenido en los últimos años (5)(13)(19). El diseño de la imagen clave a través de la simplificación de la imagen buscada ya sea por el uso de la información presente en una o varias características tales como contornos, color, texturas etc, es uno de los puntos de mayor interés ya que permite estudiar al mismo tiempo la cantidad de información necesaria para recuperar imágenes de forma eficiente (6)(7).

En este trabajo se ha desarrollado un esquema completo de almacenamiento y recuperación de información sobre una base de datos de imágenes en el que atributos tan importantes como el color o la forma son los elementos a manejar en el proceso de interacción con la base de datos. La figura 1 muestra el esquema general al que responde un modelo de recuperación de imágenes. El trabajo presentado centra su atención únicamente en el módulo de extracción de características.

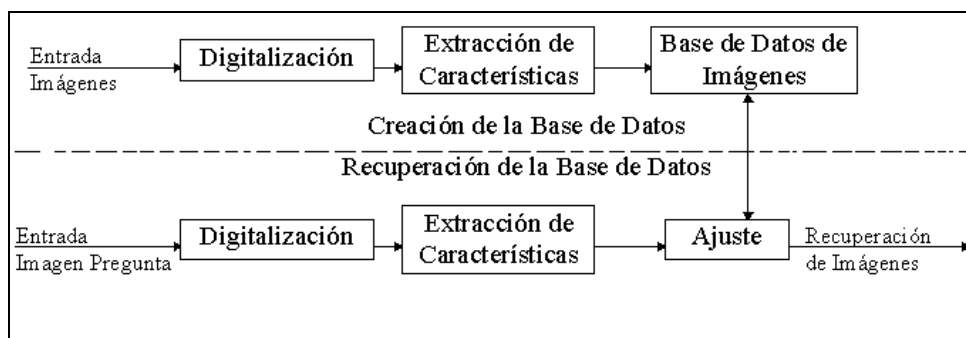


Figura 1. Esquema general de recuperación de imágenes de una base de datos

La extracción de características basadas en el contenido de la imagen, tanto en la parte de creación como de recuperación de la base de datos, trabaja a partir de imágenes en color. En la línea de determinación de los objetos presentes en la escena, la detección, simplificación y jerarquización de la información de color de los píxeles de la imagen facilita la segmentación de los mismos. Esos serán precisamente los puntos a tratar en la comunicación.

La segmentación o partición de la imagen en zonas homogéneas, entendiendo por homogeneidad la uniformidad de alguna característica, como puede ser la intensidad, el color o la textura, es uno de los primeros procesos que se aplican a una imagen en su fase de análisis o comprensión. Se aplica directamente a los píxeles de la imagen y da como resultado una serie de primitivas (regiones, contornos, ...), que aumentan el nivel de abstracción sobre nuestro conocimiento acerca del contenido de la imagen, y que usualmente son empleadas por posteriores procesos de reconocimiento o interpretación para modelizar una descripción de la escena.

El color en una escena es un atributo suficientemente discriminante para la delimitación de regiones por parte del ojo humano. Experimentos perceptuales demuestran que una persona distingue fácilmente diferencias de color aun cuando la intensidad de la imagen sea uniforme. Parece pues acertado incluir de alguna forma la información puramente cromática de una escena e integrarla con la componente de intensidad para mejorar los resultados en detección de fronteras. Sin embargo, tradicionalmente, la información de color ha sido obviada en multitud de ocasiones, bien para simplificar cálculos, bien para reducir el espacio de almacenamiento necesario para las imágenes. Trabajos recientes demuestran un creciente interés en el uso del color como atributo tanto para la segmentación de imágenes(3)(4)(17), como para su uso en la recuperación de imágenes en grandes bases de datos(18)(9).

Para adaptarse lo más posible al modo en que el ojo percibe los colores se desarrollaron los llamados *espacios perceptuales de color*, o *Espacios de Escala de Cromaticidad Uniforme (UCS)*, que tratan de describir la información de color de la forma más parecida posible a como la percibe el ser humano. En este sentido se separa totalmente la información de intensidad de la puramente cromática, de forma que para un valor constante de la primera tengamos agrupados todos los colores del espectro que poseen dicha intensidad en un único plano. Espacios de este tipo son entre otros el  $(L^*, a^*, b^*)$ , el  $(L^*, u^*, v^*)$  y el  $(H, S, I)$ .

Los enfoques que tradicionalmente se han aplicado a la segmentación de regiones en imágenes (sean en color o no) se basan en tres mecanismos diferentes: a) la detección de contornos(8)(12), b) la segmentación mediante agrupamiento de píxeles (*clustering*) (2)(21)(22)(15) y c) la segmentación por crecimiento de regiones(15)(20) como son las técnicas de *split-and-merge*(14).

De entre los tres enfoques anteriormente mencionados, el empleado en este trabajo es el de detección de contornos o fronteras seguido de un crecimiento de las regiones a partir de los mismos. Para poder buscar las variaciones locales de color en la imagen, se hará necesario definir una distancia y un orden dentro del espacio de color escogido. A partir de ellos definiremos unos operadores que nos permitirán estimar el vector gradiente en cada píxel de una imagen en color.

En la sección 2 se aborda el tema del color, y se introducen los espacios de Escala de Cromaticidad Uniforme. Las secciones 3 y 4 describen una nueva aproximación al espacio  $(H, S, I)$  (conversión desde  $(R, G, B)$ , métrica, orden y distinción entre zonas cromáticas y acromáticas), y la extensión de un algoritmo que tradicionalmente se emplea con imágenes monocromáticas. La sección 5 describe el mecanismo de simplificación y jerarquización de la información. En la sección 6 se presentan los experimentos realizados, y en la sección 7 se exponen las conclusiones.

## 2.- Espacios perceptuales de color

El color es un atributo psicológico; los espacios perceptuales describen usualmente los colores en base a su matiz, saturación e intensidad  $(H, S, I)$ . La intensidad  $I$  es una medida de la reflectancia total en la región visible del espectro, es decir, la componente acromática del color. De hecho, una imagen monocromática únicamente conserva los valores de intensidad de la imagen total. El matiz  $H$  es un atributo perceptual del color notado como rojo, amarillo, verde, azul y así sucesivamente. Estos colores se organizan en círculo en cada plano de intensidad uniforme, por lo que nuestra medida de matiz será un ángulo. La saturación es utilizada para describir la pureza de un color concreto, así un color completamente desaturado será gris, mientras que un color completamente saturado será intenso y brillante. Geométricamente, la saturación  $S$  será la distancia de la línea de grises a la que se encuentre el color en cuestión. En la figura 2 podemos apreciar la interpretación geométrica de estas tres componentes, dentro de un espacio de representación cilíndrico.

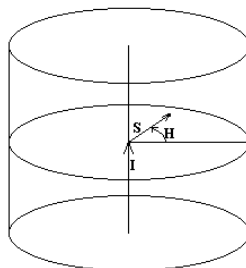


Figura 2. Interpretación geométrica de las coordenadas  $(H, S, I)$  en el espacio cilíndrico en el que se inscriben

### 3.- Aproximación al espacio (H,S,I)

El espacio de color (H,S,I) es un sistema tricolor que formaliza el sistema de color desarrollado por Munsell(10)(11). Su diseño refleja la forma en que los humanos ven el color, por lo que ofrece numerosas ventajas en procesamiento de imagen. En este trabajo se ha preferido desarrollar una aproximación propia a dicho espacio que satisfaga todas nuestras necesidades.

En el espacio (H,S,I) la componente I denota la intensidad total de brillo de un pixel dado. Las otras dos componentes contienen la información de color del pixel, formando planos circulares de intensidad constante. La componente H, o componente de matiz, representa al matiz de color utilizado y se expresa como un ángulo en el que  $0^\circ$  representa el rojo,  $120^\circ$  representa al verde y  $240^\circ$  representa al azul. Entre los  $240^\circ$  y los  $360^\circ$  se sitúan los colores no espectrales que el ojo humano percibe (Línea púrpura del diagrama de cromaticidad del espacio). La componente de saturación representa la distancia desde punto de color hasta el centro del círculo.

Geoméricamente nos aparece un espacio con forma cilíndrica, en el que en el plano de cromaticidad de una base aparece el negro, mientras que el blanco aparece en la base opuesta. Hemos de hacer notar que tanto el negro como el blanco son colores individuales, que en lugar de venir representados por puntos, ocuparían completamente cada base del cilindro, dando lugar a considerables inestabilidades en los cálculos (de hecho, el negro podría tener una distancia finita mayor que cero consigo mismo). Por esta razón se ha escogido una forma geométrica ligeramente diferente para representar el espacio de color.

En la aproximación realizada al espacio (H,S,I) en este trabajo se ha considerado una forma geométrica de dos conos unidos en su base por un cilindro (figura 3.c). El negro y el blanco puros se colocarían en los vértices de cada cono. La aparición de esta forma geométrica se debe fundamentalmente a que conforme nos movemos por el eje de intensidad hacia el blanco o el negro puros, deja de tener sentido considerar información de saturación y de matiz. El negro es la intensidad cero, y el blanco la intensidad máxima, no existe un negro (blanco) más o menos saturado, ni, por supuesto con matices de color.

#### 3.1.- Conversión entre (R,G,B) y (H,S,I)

La conversión desde el espacio (R,G,B) a un espacio (H,S,I) como el que se describe en el apartado anterior se puede llevar a cabo de la siguiente manera:

Puesto que la diagonal del cubo (R,G,B) que va de (0,0,0) a (1,1,1) es la línea de grises (figura 3.a), el primer paso consistirá en rotar el cubo RGB para que dicha diagonal quede alineada con uno de los ejes (figura 3.b). De esta forma separamos la información cromática (ejes x e y) de la que corresponde a la luminosidad (eje z):

$$x = \frac{1}{\sqrt{6}}[2R - G - B] \quad y = \frac{1}{\sqrt{2}}[G - B] \quad z = \frac{1}{\sqrt{3}}[R + G + B] \quad (1)$$

Posteriormente pasamos a coordenadas cilíndricas mediante las siguientes ecuaciones:

$$\rho = \sqrt{x^2 + y^2} \quad \phi = \arctan\left(\frac{y}{x}\right) \quad (2)$$

Estas coordenadas  $(\phi, \rho, z)$  corresponden con los valores  $(H,S,I)$ . Sin embargo, nuestro espacio sigue teniendo forma de cubo. Para convertir la parte central en un cilindro y los extremos superior e inferior en conos, dividiremos  $S$  por el valor máximo que puede tomar  $\rho$  para los valores de  $\phi$  y  $z$  dados, y lo multiplicaremos por el radio máximo correspondiente al valor de  $I$ .

La transformación queda como sigue:

$$\begin{aligned} \phi_1 &= \phi + k \frac{2\pi}{3} & , \text{ tal que } \phi_1 &\in \left[ -\frac{\pi}{3}, \frac{\pi}{3} \right] \\ \phi_2 &= \phi - \frac{\pi}{3} + k \frac{2\pi}{3} & , \text{ tal que } \phi_2 &\in \left[ -\frac{\pi}{3}, \frac{\pi}{3} \right] \\ m_1 &= \frac{\sqrt{3} - I}{\sqrt{2} \cos \phi_1} \\ m_2 &= \frac{I}{\sqrt{2} \cos \phi_2} \end{aligned} \tag{3}$$

$$S = \begin{cases} \frac{\rho \left[ 1 - 2 \left| \frac{I}{\sqrt{3}} - \frac{1}{2} \right| \right] \sqrt{3}}{\sqrt{2} \min(m_1, m_2)} & , \text{ si } I \notin \left[ \frac{1}{\sqrt{3}}, \frac{2}{\sqrt{3}} \right] \\ \frac{2\rho}{\sqrt{6} \min(m_1, m_2)} & , \text{ si } I \in \left[ \frac{1}{\sqrt{3}}, \frac{2}{\sqrt{3}} \right] \end{cases}$$

El resultado queda compuesto por los conos y el cilindro mínimos que, teniendo como eje la diagonal principal del cubo  $(R,G,B)$  original, envuelven a dicho cubo, todo ello en el interior del espacio cilíndrico que se empleará posteriormente para definir la métrica utilizada (figura 3.c).  $I$  queda en el rango  $[0, \sqrt{3}]$ ,  $H$  en  $[0, 2\pi]$ , y  $S$  en  $[0, \sqrt{2}/\sqrt{3}]$ .

En la figura 4 se muestra el espacio visto de perfil, así como un plano cromático sobre el que se sitúan el rojo, el verde y el azul.

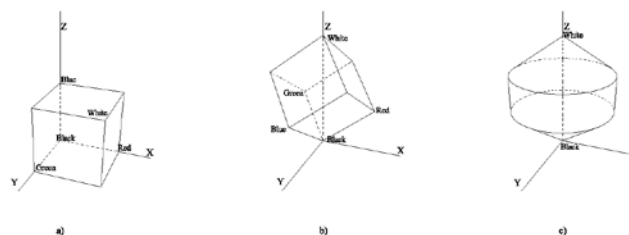


Figura 3. Representación de las transformaciones

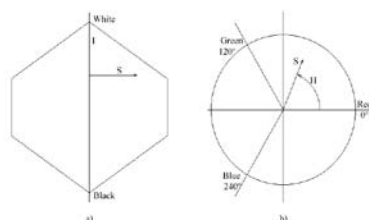


Figura 4. Proyección longitudinal del espacio  $(H,S,I)$ , b) corte transversal del mismo

### 3.2.- Areas Cromáticas y Acromáticas en el espacio $(H,S,I)$

Trabajos previos(15)(23) confirman que la componente de matiz es generalmente el atributo más discriminante del espacio (H,S,I). Cualquier método eficaz de detección de fronteras sobre imágenes en color debe tener presente su uso en regiones de información cromática. El negro, el gris o el blanco son *colores acromáticos* caracterizados únicamente por el valor de su componente de intensidad. Cuando una escena es observada bajo malas condiciones de iluminación aparecen en la imagen regiones acromáticas. Así por ejemplo, una débil iluminación provoca que el objeto adquiera un aspecto oscuro en el que no se pueden distinguir colores. De otra parte, una fuerte iluminación produce en los sensores una saturación (no confundir con la componente de saturación S de un color) y distorsión del color que inducirá un decremento del valor de saturación.

En tales condiciones, el atributo de matiz H es menos importante, llegando a ser la intensidad I el único factor discriminante en la segmentación de regiones acromáticas. De cara a la optimización del algoritmo de segmentación es necesario definir un mecanismo que nos permita clasificar los píxeles por su pertenencia a zonas cromáticas o acromáticas, para poder medir distancias de la forma más adecuada.

En nuestra aproximación hemos partido el espacio (H,S,I) definido en dos zonas delimitadas por un paraboloide centrado en el eje de intensidad (figura 5), de forma que se consideran acromáticos todos los colores del espacio que quedan dentro de dicho paraboloide, y cromáticos los que quedan fuera. A efectos de cálculo de distancias se han definido clases de equivalencia, considerando que todos los puntos del disco interior al paraboloide parametrizado situados en cada nivel de intensidad tienen un valor de saturación igual a la existente en todo su borde. Recordemos que dentro de estas zonas se supone que el ojo humano sólo distingue gris, por lo que todos esos puntos deben considerarse equivalentes independientemente de sus valores de H y S.

Para parametrizar correctamente el paraboloide, se ha realizado un experimento en el que diez personas marcaban como cromáticos o acromáticos una serie de colores seleccionados al azar por todo el espacio. Una vez representados dichos puntos, y asumiendo que la cromaticidad o acromaticidad de un color no depende de su valor de matiz (H), se ha llegado al siguiente polinomio de grado 2, que da una aproximación razonable a los resultados experimentales:

$$\Omega_1 = 0.15 + 0.2 \left( I - \frac{\sqrt{3}}{2} \right)^2 \quad (4)$$

donde  $\Omega_1$  representa el valor mínimo de saturación que debe tener un color para ser considerado como cromático, para un valor de intensidad I dado.

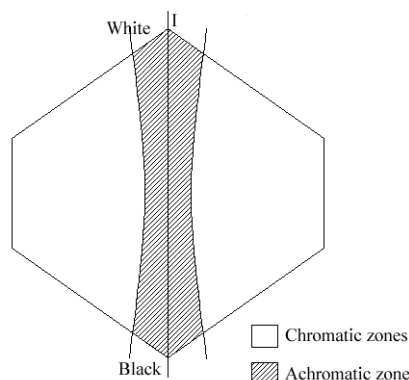


Figura 5. Representación de las zonas cromáticas y acromáticas

En los siguientes apartados se define tanto un orden como una métrica en (H,S,I). Como se verá más adelante, se ha preferido definirlos de forma que puedan aplicarse de manera

homogénea tanto a píxeles cromáticos como acromáticos. De esta forma no necesitamos distinguir explícitamente zonas cromáticas o acromáticas en la imagen, como hacen otros métodos, sino que trataremos todos los puntos por igual.

### 3.3.- Definición de una métrica en el espacio (H,S,I)

Para analizar imágenes en color es muy importante disponer de una métrica que nos permita cuantificar la diferencia (o distancia) entre dos colores. Puesto que nuestro espacio (H,S,I) está inscrito dentro de un cilindro, definiremos la distancia entre dos puntos ( $P_1, P_2$ ) cualesquiera de la siguiente forma:

$$d = ((d_I)^2 + (d_C)^2)^{1/2} \quad (5)$$

donde

$$d_I = |I_1 - I_2| \quad (6)$$

y

$$d_C = \begin{cases} ((S_1)^2 + (S_2)^2 - 2S_1S_2 \cos \phi)^{1/2} & , \text{ si } P_1 \text{ y } P_2 \text{ son cromaticos} \\ 0 & , \text{ si } P_1 \text{ y } P_2 \text{ son acromaticos} \\ S_1 - \Omega_{I_2} & , \text{ si } P_1 \text{ es cromatico y } P_2 \text{ es acromatico} \end{cases} \quad (7)$$

siendo

$$\Omega_{I_2} = \text{Umbral de cromaticidad para } I_2 \quad (8)$$

$$\phi = \begin{cases} |H_1 - H_2| & \text{si } |H_1 - H_2| > \pi \\ 2\pi - |H_1 - H_2| & \text{si } |H_1 - H_2| < \pi \end{cases} \quad (9)$$

El valor  $d_I$  es la diferencia de intensidad entre los dos puntos,  $d_C$  es la distancia entre las proyecciones de los puntos en un plano de cromaticidad, y  $\phi$  es el menor ángulo que separa  $H_1$  y  $H_2$ . En la figura 6 podemos observar la interpretación geométrica de esta métrica.

En el caso de que los dos puntos fueran acromáticos, la distancia sólo obedecería a la diferencia de intensidades. Si por el contrario sólo uno de los dos fuera acromático, se considera el punto de la clase de equivalencia de ese color más cercano a aquel con el que queremos medir la distancia, que será el que tenga el mismo valor de H y un valor de S igual al umbral de cromaticidad. La  $d_C$  se calcula entonces restándole al valor de saturación del punto cromático el umbral de cromaticidad para el nivel de intensidad del punto no cromático.

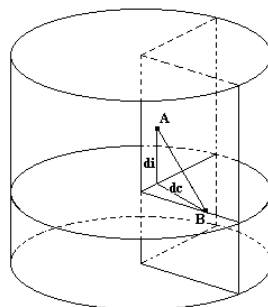


Figura 6. Representación gráfica de la distancia entre los puntos A y B en un espacio cilíndrico, siendo D la distancia,  $d_I$  la diferencia en intensidad, y  $d_C$  la distancia desde B hasta la proyección de A sobre su plano de cromaticidad

### 3.4.- Especificación de un orden en el espacio (H,S,I)

Para completar nuestro algoritmo, como se verá más adelante, necesitaremos definir un orden dentro del espacio (H,S,I) que nos permita, dados dos puntos, decidir cuál de los dos se considera mayor.

Puesto que no existe ninguna forma natural universalmente aceptada de ordenar los colores, definiremos nuestro orden de manera arbitraria. Una técnica simple y que da buenos resultados en la práctica consistiría en evaluar una de las componentes en los dos elementos a ordenar, y elegir como elemento mayor el que posea un valor más grande en dicha componente. En caso de igualdad pasaríamos a la siguiente componente, y así sucesivamente.

En nuestro caso, dados dos píxeles  $(H_1, S_1, I_1)$  y  $(H_2, S_2, I_2)$ , comparamos en primer lugar los valores de I, luego los de S, y finalmente los de H. Para comparar esta última componente el criterio empleado es el siguiente: si el ángulo de menor módulo para pasar de  $H_1$  a  $H_2$  es positivo, elegimos  $(H_2, S_2, I_2)$  como elemento mayor, y en caso contrario nos quedamos con  $(H_1, S_1, I_1)$ .

Si los dos puntos que queremos ordenar fueran acromáticos, sólo atenderemos a su diferencia de intensidades. En el caso de que sólo uno fuera acromático, a igualdad de intensidad se considera mayor el punto cromático (que necesariamente tiene mayor saturación).

## 4.- Detección de Fronteras en imágenes de color

El enfoque de segmentación basado en detección de contornos se apoya en los llamados operadores diferenciales, que sirven para detectar discontinuidades o cambios locales en los valores de los píxeles.

En este apartado generalizaremos las técnicas de detección de fronteras para imágenes monocromáticas a las definidas en nuestro espacio de color. La obtención de la estimación del gradiente en espacios de color se hará a partir de la distancia y el orden definidos para dicho espacio.

### 4.1.- Estimación del vector gradiente en imágenes en color

Sea  $f(x,y)$  una función escalar definida sobre  $\mathfrak{R}^2$ , como por ejemplo una imagen monocromática. Se define el gradiente de  $f$ , notado como  $\nabla f$ , de la siguiente forma:

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} \quad (10)$$

De esta definición se extrae que el gradiente de  $f$  es un vector de  $\mathfrak{R}^2$ , orientado en la dirección donde la variación de  $f$  es máxima. El módulo de dicho vector es la magnitud de la derivada de  $f$  en la dirección del propio vector.

La anterior definición del concepto de gradiente no se aplica directamente a imágenes multibanda, puesto que no existe el concepto de derivada parcial de una función cuyos valores están en  $\mathfrak{R}^3$ . Para tratar de remediar este problema, surgen diferentes aproximaciones de estimación de lo que se ha dado en llamar *gradiente multispectral*(16).

Muchas de estas aproximaciones se basan en la combinación (usualmente lineal) de los



gradientes individuales de cada banda de la imagen. Si  $P(x,y)$  es el color del punto  $(x,y)$  y  $(P_1, P_2, P_3)$  sus componentes en el espacio de color:

$$\nabla C = F(\nabla P_1, \nabla P_2, \nabla P_3) \quad (11)$$

donde  $\nabla C$  es el gradiente en el punto  $(x,y)$ ,  $F$  representa cualquier función y  $\nabla P_1, \nabla P_2, \nabla P_3$  son los gradientes individuales en  $(x,y)$  de cada una de las componentes de  $P$ . Es decir, se trata la imagen como si fueran tres monocromáticas, y una vez obtenidos los gradientes de cada imagen, se busca una manera de combinarlos. El problema de este método es precisamente la artificiosidad de esa combinación.

Otras técnicas tratan de estimar el módulo de las componentes del vector gradiente en función de alguna distancia definida sobre el espacio de color. Puesto que la derivada en una dirección se puede aproximar como la diferencia entre un píxel y el siguiente en esa misma dirección, podemos usar la distancia en un espacio de color para estimar el módulo de dicha diferencia:

$$|\nabla C_x| = \left\| \frac{\partial P}{\partial x} \right\| \cong d(P_{x,y}, P_{x+1,y}) \quad |\nabla C_y| = \left\| \frac{\partial P}{\partial y} \right\| \cong d(P_{x,y}, P_{x,y+1}) \quad (12)$$

En el presente trabajo se ha preferido desarrollar una aproximación parecida, expresando la derivada de los valores de la imagen como una distancia, si bien como punto novedoso se ha añadido el concepto de orden, lo que nos permitirá estimar no sólo el módulo de las componentes del vector gradiente, sino también su dirección. Orientaremos pues la componente obtenida del vector gradiente en la dirección del mayor valor de píxel según el orden que tenemos definido

En lo que respecta al orden en sí, baste decir que aunque es arbitrario sigue siendo coherente con el orden natural en la escala de grises, por lo que en la práctica no produce inestabilidades detectables, dando excelentes resultados experimentales.

## 4.2.- Extensión del Filtro de Canny a imágenes en color

La información sobre el vector gradiente (magnitud y dirección) que proporciona el operador de primer orden empleado es fundamental en la interpretación de las fronteras. En esta sección extenderemos el uso de máscaras de primer orden basadas en la Gaussiana, como hace el filtro de Canny(1), de imágenes monocromas a imágenes en color.

La convolución de una imagen de niveles de gris con una primera derivada parcial de Gaussiana se puede llevar a cabo convolucionando sucesivamente la imagen con dos máscaras unidimensionales perpendiculares, siendo la primera la derivada parcial de la Gaussiana y la segunda la propia Gaussiana. Puesto que el resultado de esta primera convolución ya serían números reales (positivos o negativos según la dirección del vector gradiente), nuestro problema se reduce entonces a una sólo dimensión, pero con máscaras de longitud arbitraria.

Para explicar esto haremos uso de la siguiente propiedad:

$$\frac{\partial G(x, y)}{\partial x} = - \frac{\partial G(-x, y)}{\partial x} \quad (13)$$

Luego, en el caso discreto, la acción de aplicar una máscara de derivada parcial en  $x$  de

Gaussiana se puede expresar de la siguiente manera:

$$\sum_k m_k x_{i+k} - m_k x_{i-k} \quad (14)$$

Llegados a este punto parece lógico pensar que podemos emparejar los píxeles que tienen igual peso (están afectados por el mismo coeficiente). Agrupando entonces los términos y emparejando los correspondientes a píxeles equidistantes del punto donde la máscara se encuentra centrada llegamos a la siguiente expresión:

$$\sum_k m_k (x_{i+k} - x_{i-k}) = \sum_k m_k c(x_{i+k}, x_{i-k}) d(x_{i+k}, x_{i-k}) \quad (15)$$

siendo  $d(x_i, x_j)$  la distancia entre los colores de los píxeles  $x_i$  y  $x_j$ , y

$$c(x_i, x_j) = \begin{cases} 1 & \text{si } x_i > x_j \\ -1 & \text{si } x_i < x_j \end{cases} \quad (16)$$

una función de orden como la definida en el apartado anterior.

Como se puede apreciar, basta con definir una métrica y un orden en el espacio de color para poder convolucionar la imagen con la derivada parcial en una dirección, y posteriormente convolucionar perpendicularmente el resultado (recordemos que ya son números reales) con una máscara de Gaussiana. A partir de este punto se aplican los restantes pasos del algoritmo de Canny, es decir, la supresión no maximal y el paso de histéresis para terminar de aislar los puntos frontera más prometedores.

Con este resultado se ha generalizado el filtro de Canny de detección de bordes al caso de imágenes en color. Se puede observar cómo además del uso de la intensidad como información para detectar bordes, se está empleando información proveniente del plano de color y saturación. Es fácil apreciar, además, que este filtro aplicado a imágenes en tonos de gris es equivalente al filtro de Canny, dando por supuesto el mismo resultado.

## 5.- Segmentación de Imágenes

La segmentación de imágenes es, en tratamiento de señales, uno de los primeros procesos de *condensación de la información*. Se aplica directamente a los píxeles de la imagen y produce un reagrupamiento de los mismos de acuerdo a algún criterio para la formación de contornos, regiones, etc. Esas primitivas describen y modelizan la imagen y son, en general, utilizadas por los procesos posteriores de reconocimiento/interpretación de las escenas.

De forma general, como veíamos en la introducción, los algoritmos de segmentación de imágenes (sea cual sea el tipo de imagen: color, monocroma...) pueden ser clasificados en tres grandes categorías, atendiendo a su forma de actuación.

Uno de las principales objeciones al uso de técnicas de crecimiento de regiones se deriva de la necesidad del método de seleccionar un píxel de la imagen, la semilla, alrededor del cual crecerá la región.

En la aproximación realizada en este trabajo se ha optado por una técnica de crecimiento de regiones a partir de la información de gradiente obtenida en la detección de contornos.

### 5.1.- Crecimiento de Regiones

## Sembrado de semillas

Como anteriormente mencionábamos, el primer paso a la hora de llevar a cabo una técnica de crecimiento de regiones es la elección de los puntos semilla a partir de los cuales se irán extendiendo las diferentes regiones dentro de la imagen. La técnica de crecimiento aquí propuesta trata de sacar partido de la información de gradiente y fronteras que de la imagen se tiene, haciendo que las regiones resultantes no rebasen las fronteras obtenidas en el paso previo, y que éstas crezcan además en función del gradiente de color de la imagen previamente calculado. De esta forma conseguimos que las propias fronteras actúen de una parte como muros de contención del proceso de inundación, y de otra que sean ellas, más concretamente la información de gradientes conseguida, quienes den el criterio de similitud en el proceso de agregación de píxeles a las distintas semillas.

Para evitar que las regiones rebasen las fronteras ya calculadas sembraremos para cada segmento de frontera a lo largo de todo su recorrido dos semillas diferentes, a derecha e izquierda (coincidiendo con la propia frontera). De esta manera lo primero que definimos es una serie de puntos de unión de las diferentes regiones, coincidentes con las fronteras previamente calculadas en pasos anteriores.

Con el propósito de suprimir inconsistencias en el sembrado y simplificar el algoritmo, eliminaremos todos los puntos triples del mapa de fronteras, de forma que todos los segmentos que nos queden carezcan de ramificaciones. Un punto triple (ver figura 7.b) es aquel que tiene tres vecinos (recordemos que después del proceso utilizado de *thinning*, los puntos con valor 1 sólo pueden tener 0, 1, 2 ó 3 vecinos a 1). Finalmente, marcaremos como semillas todos los puntos que queden inmediatamente a la derecha e izquierda del segmento de frontera, empleando una etiqueta para los de la derecha y otra para los de la izquierda. En la figura 7.a se muestra dicho proceso de siembra a lo largo de un segmento de frontera detectado.

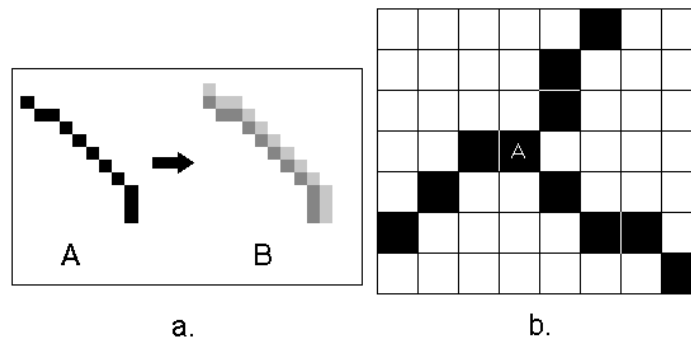


Figura 7. a) Muestra el proceso de siembra con dos etiquetas distintas a un lado y sobre la frontera, b) muestra un ejemplo de punto triple.

Esta aproximación, de una parte garantiza que las regiones resultantes no se invadirán unas a otras, y de otra, permite que segmentos no conectados en el proceso de detección de fronteras que definen una forma quedarán sembrados por semillas diferentes. Esto dará lugar a un recubrimiento de la forma por regiones de características (como el color) muy similares entre sí (Ver figura 8), lo que permitirá agrupar esas regiones en una única atendiendo a la información de vecindad que se tiene.

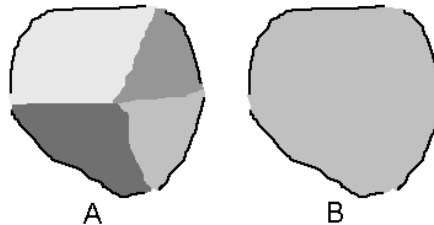


Figura 8. a) Muestra las regiones generadas tras el sembrado inicial y su crecimiento, b) muestra la forma obtenida mediante el agrupamiento de a.

## Crecimiento de regiones

Nuestro algoritmo de crecimiento de regiones se basa en un descenso controlado a través de las pendientes de la imagen de gradientes. Recordemos que nuestras semillas están colocadas precisamente en las fronteras de la imagen, que corresponden con los máximos (o crestas) que presenta la imagen de gradientes.

El algoritmo de propagación es simple: para cada píxel se miran los vecinos no etiquetados, y se etiqueta con el mismo valor del píxel original aquel que tenga un valor de magnitud de gradiente más cercano al que poseía el píxel originalmente.

Este mecanismo de crecimiento por sí solo es insuficiente, puesto que hace crecer las regiones de manera homogénea, independientemente de las diferencias reales de magnitud de gradiente. De esta forma una región que presentara una pronunciada pendiente en la imagen de magnitudes de gradiente crecería igual que otra con poca o ninguna pendiente. Para evitar esto se limita el valor mínimo  $M$  de magnitud de gradiente que puede tener un píxel para que la frontera se propague. Cuando el paso de propagación no hace ningún cambio, se decrementa el valor de  $M$  y se sigue propagando. El algoritmo de crecimiento queda, pues, como sigue :

### ALGORITMO:

```
G(x,y) : Imagen de magnitud de gradiente.  
E(x,y) : Imagen de etiquetas  
M=max(G)  
m=min(G)  
k=M  
Repetir mientras k>=m  
  Repetir mientras haya cambios en E(x,y)  
    Para todo E(x,y), sea E(x',y') un vecino no etiquetado  
    E(x',y')=E(x,y) sii |G(x,y)-G(x',y')| es mínimo y G(x',y')>=k  
  Fin Repetir  
  k=k-Δk  
Fin Repetir
```

## Interpolación de las regiones

Una vez que tenemos extendidas nuestras regiones podemos efectuar una interpolación simple de los valores (H,S,I) que nos permita una mejor visualización de la información de dependencias contextuales de los parches o regiones calculados.

Para poder realizar la interpolación calcularemos la media de los valores (H,S,I) de la imagen original en toda la región. El mecanismo empleado para calcular dicho promedio es el siguiente:

$$x = \frac{\sum_{i=1}^n S_i \cos(H_i)}{n} \quad y = \frac{\sum_{i=1}^n S_i \sin(H_i)}{n} \quad (17)$$

$$\bar{H} = \arctan\left(\frac{y}{x}\right) \quad \bar{S} = \sqrt{x^2 + y^2} \quad \bar{I} = \frac{\sum_{i=1}^n I_i}{n} \quad (18)$$

### 5.2.- Partición final de las regiones

Puesto que nuestro sembrado inicial partía de los segmentos de frontera calculados inicialmente, y dado que estos segmentos pueden resultar incompletos, existe el riesgo de que alguna región se extienda a través de algún hueco en las fronteras, invadiendo lo que correspondería a otra región (Ver figura 9a). También es posible que tengamos un segmento de frontera que delimita por uno se sus lados dos o más regiones diferentes, por lo que estaríamos cometiendo un error en el sembrado (Ver figura 9b).

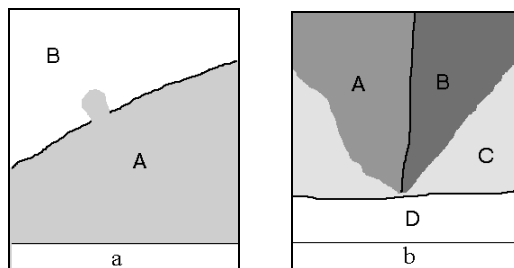


Figura 9. Casos de error que se pueden producir en el crecimiento de las regiones

Para solucionar estos posibles problemas se particionan algunas regiones. El criterio usado para determinar si una región debe particionarse o no atiende a que la media de las distancias (H,S,I) entre los valores interpolados y los de los píxeles originales supere o no un umbral dado.

El mecanismo de partición diseñado divide las regiones por su parte más estrecha, buscando los dos puntos del contorno cuyo cociente entre distancia euclídea y distancia más corta a través del contorno sea mínima, bajo la restricción de que el segmento que los una pertenezca a la región.

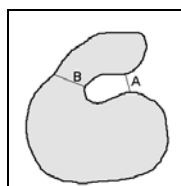


Figura 10. El segmento A presenta un mínimo, pero no es interior a la región. El corte se hará por B

## 6.- Experimentos realizados

Para llevar a cabo los experimentos se han seleccionado imágenes de 256x256 píxeles, con un byte para cada banda y píxel. Algunas de ellas se han sintetizado expresamente para remarcar las situaciones en las que el filtro de Canny puede presentar problemas.

La implementación de todos los filtros y operaciones necesarios para la realización de los experimentos se ha efectuado bajo Khoros 2.1. Con este propósito se ha diseñado un *Toolbox* con las diversas transformaciones entre espacios de color, así como las librerías necesarias para calcular las distancias dentro del espacio (H,S,I), para llevar a cabo todos los pasos de cada filtro y el crecimiento de las regiones a partir de los segmentos de frontera detectados.

### 6.1.- Resultados obtenidos

Contenido de las figuras de resultados:

1. En (a) se muestran las imágenes originales
2. En (b) el resultado de aplicar el proceso de crecimiento de regiones con la posterior partición e interpolación de las mismas
3. En (c) el resultado de aplicar el filtro de Canny a la media aritmética de las tres bandas de color con los siguientes parámetros:  $\sigma = 1$ , umbral superior de histéresis  $h = 0.9$ , y umbral inferior  $l = 0.5$ .
4. En (d) el resultado obtenido al aplicar el filtro de color tipo Canny a cada imagen, con los mismos parámetros  $\sigma$ ,  $l$  y  $h$ . El paraboloide utilizado para delimitar las zonas cromáticas y acromáticas es la revolución de la parábola dada en [4].

Todas las imágenes de resultados han sido tratadas por un proceso de *thinning* para que las fronteras que aparecen en las imágenes tengan de grosor un píxel.

La figura 11 corresponde a una imagen sintética que representa una casa. La mayor parte de los colores que se han escogido van al mismo tono de gris al efectuar la media de las tres bandas. Es por esta razón que el filtro de Canny es incapaz de encontrar estas fronteras, mientras que los filtros basados en la distancia (H,S,I) detectan la diferencia de color entre las regiones sin dificultad.

En la figura 12, correspondiente a la imagen del escorpión, se puede observar una característica importante: el cefalotórax del escorpión posee un dibujo longitudinal que pasa totalmente desapercibido para Canny, mientras que los otros tres filtros lo detectan.

La figura 13, que representa a una rana verde sobre una hoja roja, presenta un fuerte contraste para nosotros, contraste que queda muy mermado cuando se pasa a tonos de gris. Esta es la razón por la cual el filtro de Canny no es capaz de capturar todo el contorno de la rana. En la parte inferior derecha podemos ver cómo la textura que presenta la imagen es tratada por los filtros tipo Canny de una manera más coherente, mientras que los otros dos devuelven fronteras que parecen aleatorias.

La figura 14, que representa a un tucán, pone de manifiesto un fenómeno parecido. En la base del pico de este ejemplar se puede apreciar una diferencia de color de amarillo a azul, mientras que en la imagen de grises tenemos dos tonos muy similares. El filtro de Canny no detecta esta frontera.

En todas ellas se puede observar que la interpolación de los parches muestra visualmente como todos los objetos y rasgos importantes de información de cara a la interpretación de la escena han sido recogidos.

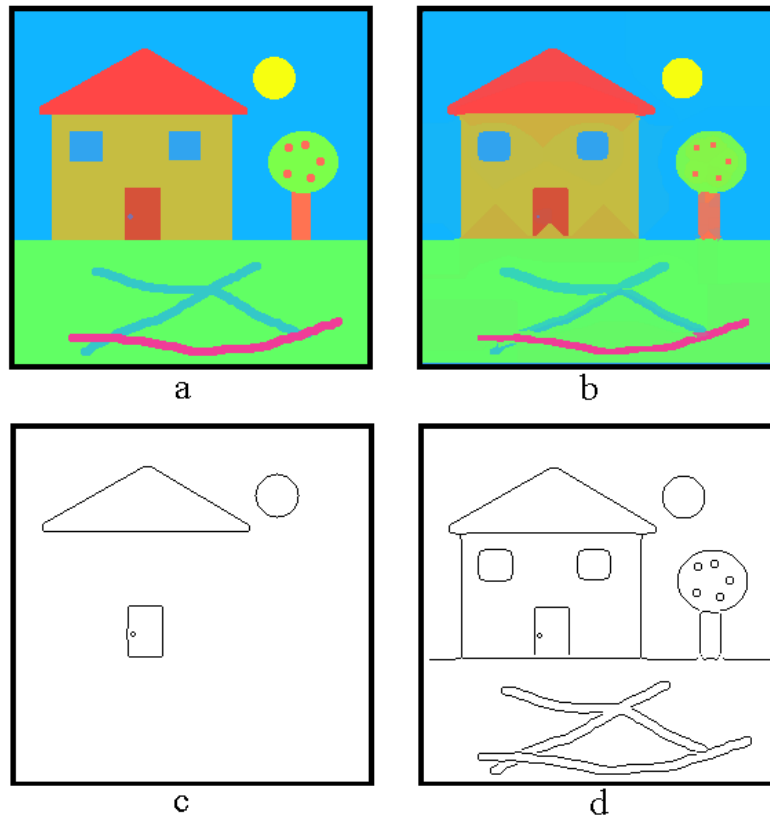


Figura 11. a) Imagen original, b) imagen de regiones interpoladas, c) segmentación de Canny sobre niveles de gris, d) segmentación de Canny sobre (H,S,I)

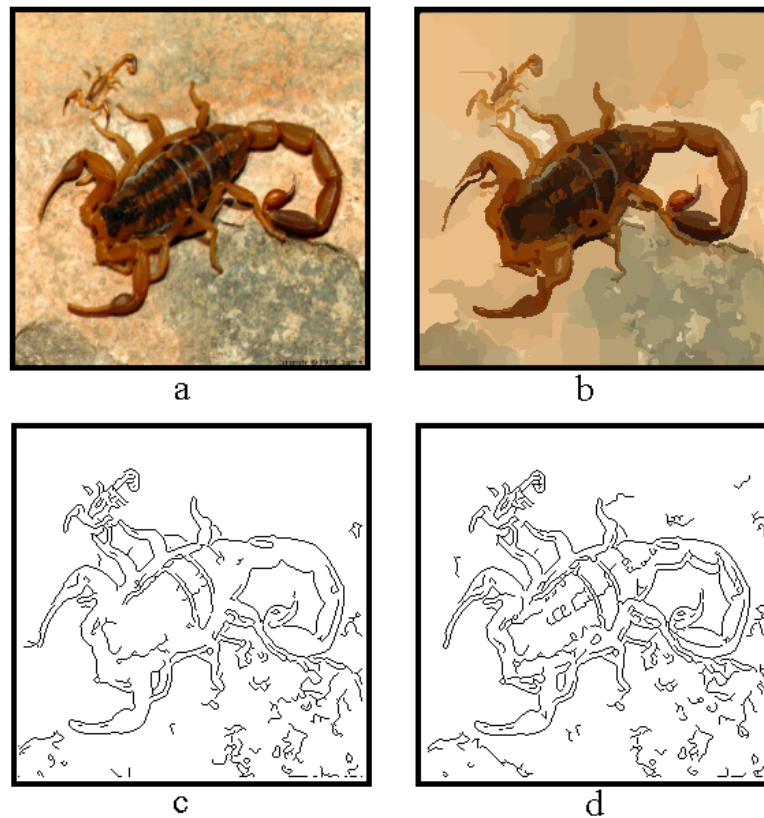


Figura 12. a) Imagen original, b) imagen de regiones interpoladas, c) segmentación de Canny sobre niveles de gris, d) segmentación de Canny sobre (H,S,I)

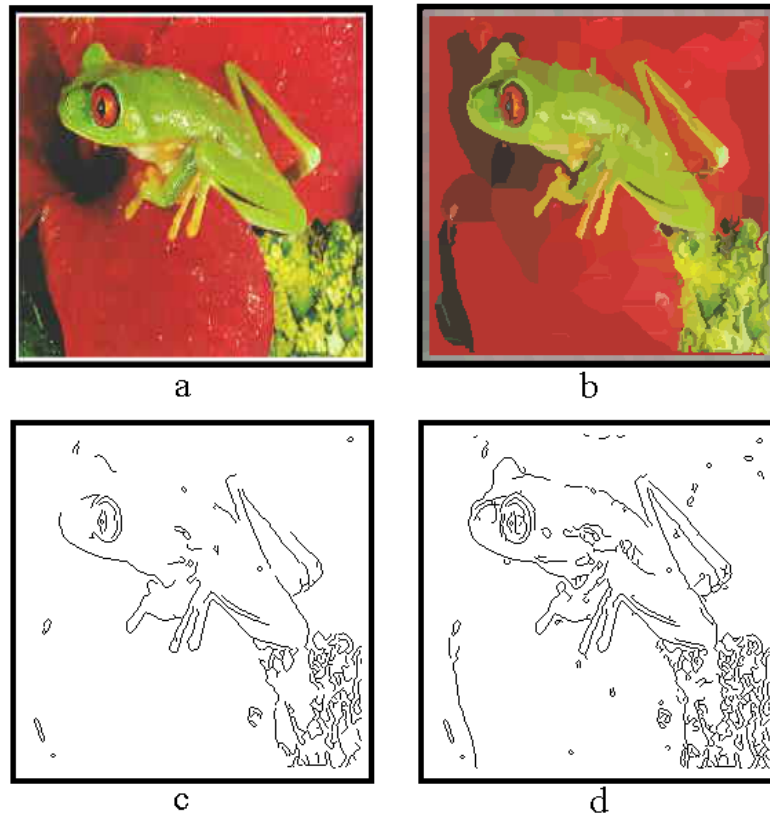


Figura 13. a) Imagen original, b) imagen de regiones interpoladas, c) segmentación de Canny sobre niveles de gris, d) segmentación de Canny sobre (H,S,I)

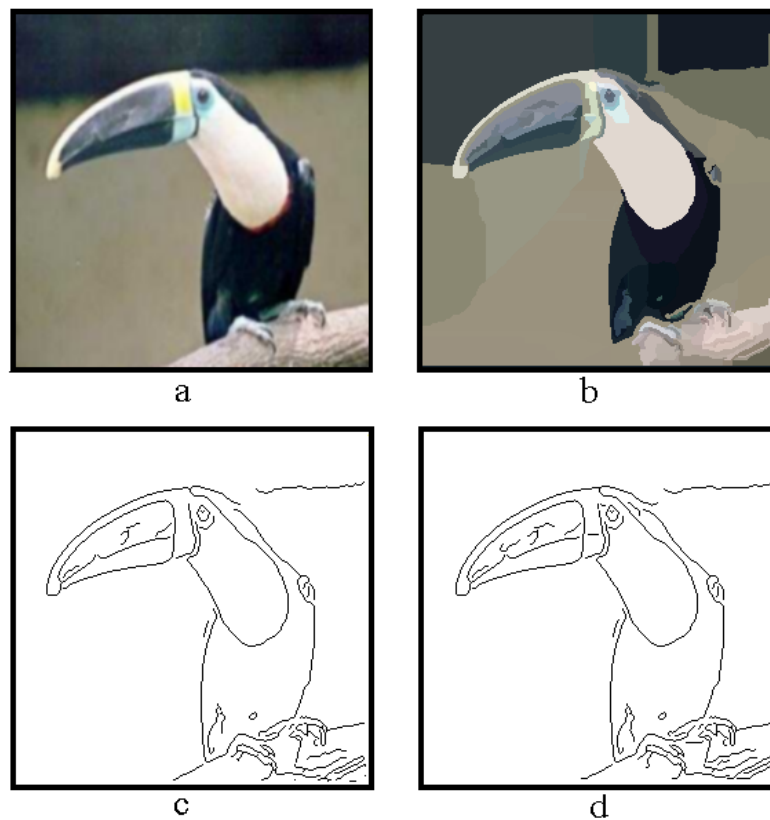


Figura 14. a) Imagen original, b) imagen de regiones interpoladas, c) segmentación de Canny sobre niveles de gris, d) segmentación de Canny sobre (H,S,I)



## 7.- Conclusiones

Quedan demostrados tanto el interés como la gran utilidad que tiene emplear espacios de color de Escala de Cromaticidad Uniforme a la hora de detectar fronteras en una imagen, en lugar de reducir ésta a niveles de gris, (pérdida de una fuerte componente perceptual de información). Se ha descrito una nueva aproximación al espacio de color (H,S,I), con una métrica propia y un mecanismo para separar las zonas cromáticas y acromáticas del espacio.

En líneas generales podemos decir sobre el método diseñado que::

- Se separa de forma natural la información de intensidad de la información estrictamente cromática.
- La especificación de los colores en este tipo de espacios es más cercana a la que proporciona la intuición humana.
- Las diferencias perceptuales entre colores pueden ser representadas directamente con la métrica del espacio.
- Podemos determinar fácilmente qué puntos deben ser tratados como cromáticos, y qué puntos deben ser tratados como grises.
- Dentro del esquema general de recuperación de imágenes en una base de datos se inserta un módulo de recubrimiento de regiones capaz de estructurar la información de las mismas en un orden de relación superior que posibilite con las técnicas adecuadas, actualmente en desarrollo, la aparición de las formas componentes de las escenas a analizar.

Empleando la definición de orden y métrica dada para este espacio (H,S,I), se ha desarrollado una extensión del operador multiescala de Canny, consiguiendo como resultado una técnica bastante robusta para la detección de fronteras en imágenes en color, y dejando abierto un claro camino para la aplicación de técnicas multiresolución de forma directa a imágenes multibanda.

Uno de los resultados más interesantes que se desprenden de este trabajo es que podemos aplicar operadores del *tipo gradiente* (se dice aquí *tipo gradiente* porque el concepto de gradiente como tal no existe en espacios de representación de más de una dimensión) a imágenes multibanda con sólo definir una métrica y un orden en el espacio de representación.

El mayor inconveniente es quizás la necesidad de definir ese orden, puesto que no existe uno universalmente aceptado en estos espacios. El orden natural en  $\Re$  es evidente, pero la cosa no es tan simple en  $\Re^2$  y espacios de orden superior. Esto nos lleva a tener que definir nuestro orden de forma arbitraria, corriendo el riesgo de que cause problemas a la hora de aplicar el operador.

Cuatro aportaciones fundamentales han sido puestas de manifiesto en el presente trabajo:

- Utilización de la información perceptual que el color aporta en el esquema de detección de fronteras diseñado.
- Extensión del algoritmo detector de fronteras de Canny a imágenes en color.
- Crecimiento de las regiones de recubrimiento de la imagen a partir de la información de gradiente aportada por el algoritmo detector de fronteras.
- Adecuación de la información de regiones componentes de la imagen en base a una estructura jerárquica de dependencias de vecindad entre las mismas.

## Bibliografía

1. Canny, 1986, *A computational approach to edge detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, pp. 679-698
2. M. Celenk, 1990, *A Color Clustering Technique for Image Segmentation*, Computer Vision, Graphics and Image Processing, num. 52, pp.145-170
3. M. Chapron, 1992, *A new chromatic edge detector used for color image segmentation*, 11th IAPR International Conference on Pattern Recognition, The Hague, vol. 3, pp. 311-314
4. G. Healey, 1992, *Segmenting images using normalize color*, IEEE Trans. Syst. Man. Cybern., vol. 22, pp. 64-73
5. Charles E. Jacobs; Adam Finkelstein; David H. Salesin, 1995, *Fast multiresolution image quering*, Proceeding of SIGGRAPH 95, pp. 277-286
6. Anil K. Jain; Aditya Vailaya, 1996, *Image retrieval using color and shape*, Pattern Recognition. Vol. 29, Num. 8, pp.1233-1244
7. Mohan S. Kankanhalli; Babu M. Mehtre; Jian Kankg Wu, 1996, *Cluster-based color matching for image retrieval*, Pattern Recognition. Vol. 29, Num. 4, pp.. 701-708
8. R. Machuca, K. Phillips, 1983, *Application of vector fields to image processing*, IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 5, pp. 316-329
9. Babu M. Mehtre; Mohan S. Kankanhalli; A. Desai Narasimhalu; Guo Chang Man, 1995, *Color matching for image retrieval*, Pattern Recognition Letters. Num 16, pp. 325-331
10. A.H. Munsell, 1913, *Munsell book of color*, Munsell color Co., 2441 North Calvert St., Baltimore, Md
11. A.H. Munsell, 1939, *A color notation (8th ed.)*, Munsell color Co., Boston
12. R. Nevatia, 1976, *A color edge detector*, 3rd International Joint Conference on Pattern Recognition, pp. 829-832
13. W. Niblack, R. Barber, W. Equitz, M. Flickner. E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, G. Taubin, 1993, *The QBIC project: Quering images by content using color, texture and shape*, Storage and Retrieval for Image and Video Databases, Vol. 1908, Proceeding of the SPIE, pp. 173-187
14. R. Ohlander, K. Price and D.R. Reddy, 1978, *Picture segmentation using a recursive region splitting method*, Comput. Graphics and Image Processing, vol. 8, pp. 313-333
15. P. Pujas, M.J. Aldon, 1994, *Segmentation des images couleur: une méthode perceptuelle*, In 3rd International Conference to Real and Virtual Worlds, pp. 1-12
16. P. Pujas, 1996, *Analyse d'images couleur et fusion d'images 3D et couleur*, Tesis Doctoral, Universidad de Montpellier
17. R. Schettini, 1993, *A segmentation algorithm for color images*, Pattern Recognition Letters, vol. 14, pp. 499-506
18. M.J. Swain, D.H. Ballard, 1991, *Color indexing*, International Journal of Computer Vision, vol. 7:1, pp. 11-32
19. M.J. Swain, 1993, *Interactive indexing into image database*. Storage and Retrieval for Image and Video Databases, Vol. 1908, Proceeding of the SPIE, pp. 95-103
20. R. Taylor, P. Lewis, 1992, *Color image segmentation using boundary relaxation*, 11th IAPR International Conference on Pattern Recognition, The Hague, vol. 3, pp. 72-1-724
21. S. Tominaga, 1986, *Color image segmentation using three perceptual attributes*, CVPR, pp. 628-630
22. S. Tominaga, 1992, *Color classification of natural color images*. Color Res. Appl. vol. 17, pp. 230-239
23. D. Tseng, C. Chang, 1992, *Color segmentation using perceptual attributes*, 11th IAPR International Conference on Pattern Recognition, The Hague, vol. 3, pp. 228-231