

III CAPÍTULO

SEGMENTACIÓN DE IMÁGENES EN COLOR

3.1. Introducción

El primer paso en el análisis de imágenes consiste generalmente en segmentar la imagen. La *segmentación* subdivide la imagen en sus partes constituyentes, que se denominan regiones o segmentos. El nivel al que se lleva a cabo esta subdivisión depende del problema a resolver. Esto es, la segmentación deberá detenerse cuando los objetos de interés de una aplicación hayan sido aislados. En general, la segmentación autónoma es una de las tareas más difíciles del procesamiento de imágenes y, a su vez, determina el éxito o el fracaso del análisis. Por esta razón, se debe poner un considerable cuidado en aumentar la probabilidad de tener una segmentación robusta [González, 1996].

A lo largo de los años se ha centrado mucha más atención a la segmentación de imágenes en escala de grises que de imágenes en color [Acha, 2002].

La segmentación de imágenes monocromáticas con respecto a la segmentación de imágenes en color es una tarea con diferencia más sencilla. Generalmente se basan en una de las dos propiedades básicas de los valores del nivel de gris: discontinuidad y similitud. En la primera categoría, el método consiste en dividir una imagen basándose en los cambios bruscos de nivel de gris. Las principales áreas de interés de esta categoría son la detección de puntos aislados y la detección de líneas y bordes de una imagen. Los principales métodos de la segunda categoría están basados en la umbralización, crecimiento de región y división y fusión de regiones [González, 1996].

Son las nuevas aplicaciones, como las bases de datos multimedia, fotografías digitales y procesamiento de datos visuales de Internet, lo que ha hecho que aparezca un nuevo interés en la segmentación, y en particular, en la de las imágenes en color [Acha, 2002].

Existe un gran número de algoritmos de segmentación distintos. En este trabajo se han utilizado distintas técnicas que tratan de resolver el difícil problema de segmentar las quemaduras de la piel sana, donde el color juega un papel principal. Los diferentes métodos de segmentación mezclan técnicas básicas de procesamiento de imágenes como son el filtrado, etc. A lo largo de este capítulo se van a exponer dichas herramientas, que se entrelazarán en distintos órdenes, dando lugar a distintos algoritmos de segmentación.

3.2. Técnicas de procesamiento inicial de la imagen

3.2.1. Cuantificación del color

En este trabajo no se utilizará esta técnica por ello no la comentaremos aquí, tan solo una breve referencia.

El objetivo de la *cuantificación de color* [Jain, 1972] [Domański, 1998] es realizar una conversión entre una imagen de color real y una imagen definida por una paleta de colores (imagen indexada). No trabajaremos con esta técnica al haberse demostrado en estudios anteriores que se obtenían mejores resultados sin su utilización [Acha, 2002].

3.2.2. Filtrado de imágenes

El filtrado es uno de los primeros pasos en la mayoría de las aplicaciones de análisis de imágenes y, desde un punto de vista general, es usado para eliminar información no útil que puede distorsionar cualquier procesamiento posterior, es decir, lo que se conoce como la componente de ruido [Schmid, 1999].

Desde el punto de vista de la segmentación, la eliminación de ruido mediante filtrado debería conservar, por otra parte, los contornos de los objetos que se quieren separar lo mejor posible. Dependiendo del tipo de filtro que se escoja, esto se consigue con una mayor o menor precisión. En los apartados siguientes se desarrollarán los fundamentos de los principales filtros usados para eliminar ruido y preparar la imagen para una posterior segmentación. Se concluirá que el filtrado de difusión anisótropa es el que más se adecua a nuestras necesidades porque, en el caso ideal, este procedimiento converge a una imagen en la que sólo se mantienen los bordes relevantes y donde el ruido y las pequeñas estructuras son eliminadas.

3.2.2.1. Filtro gaussiano

La respuesta impulsiva de un filtro gaussiano unidimensional se basa en la expresión de una *distribución gaussiana o normal*, que se define como [Haykin, 1989]

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{\pi x^2}{2\sigma^2}\right\} \quad \text{Ecuación III-1}$$

donde σ es un parámetro de diseño.

La respuesta en frecuencia del filtro sigue siendo una gaussiana y viene dada por

$$G(\omega) = \exp\left\{-\frac{\omega^2}{2\sigma_\omega^2}\right\} \quad \text{Ecuación III-2}$$

donde $\sigma_\omega = 1/\sigma$.

El filtro es utilizado para suavizar la señal según el parámetro σ_ω . Para el caso bidimensional el ruido, las pequeñas estructuras y los bordes de los objetos son afectados de la misma manera: las componentes geométricas y espaciales son modificadas y los bordes y ángulos se suavizan [Weeks, 1996].

Desventajas del filtro gaussiano:

No preserva los bordes en la imagen filtrada, aunque sus propiedades de eliminación de ruido son bastante buenas. Se utilizan mucho en el análisis escala-espacio y en la representación piramidal (multirresolución) de una imagen [Burt, 1983]. Por los inconvenientes que presenta, este tipo de filtro no será utilizado en este trabajo.

3.2.2.2. Filtro de mediana

El *filtro de mediana* se define como aquél en el que cada píxel se sustituye por el valor de mediana calculado en un entorno del píxel. El tamaño del filtro define cómo de abruptos deben ser los bordes para que se preserven.

Cuando el objetivo es más la reducción del ruido que el difuminado, el uso de filtros de mediana representa una posibilidad alternativa, especialmente, cuando el ruido es de tipo *sal y pimienta*, es decir, pequeños puntos con valores muy diferentes entre sí distribuidos por toda la imagen [Acha, 2002].

Este tipo de filtro no será utilizado en este trabajo, por no adecuarse a nuestras necesidades.

3.2.2.3. Filtro de difusión

El filtrado de difusión suaviza la imagen a la par que mantiene la información de los bordes, lo cual es muy deseable en aplicaciones de segmentación. Se le ha de prestar especial atención por su aplicación posterior en los algoritmos propuestos en este trabajo.

Un ejemplo de difusión es la transferencia de calor dentro de la materia [Haberman, 1987]. Para expresar este fenómeno físico se utiliza la ecuación diferencial en derivadas parciales (PDE, siglas en inglés de *Partial Differential Equation*) de tipo parabólico

$$\frac{\partial T}{\partial t} = \text{div}(k\nabla T) + Q \quad \text{Ecuación III-3}$$

- ∇ es el operador *gradiente*
- *div* representa al operador *divergencia*
- T es la temperatura del medio
- k es la *difusividad térmica* o *coeficiente de difusión*
- Q representa la energía calorífica generada por posibles fuentes internas de calor.

La formulación básica se obtiene cuando el medio se supone homogéneo, sin fuentes y con una difusividad térmica constante.

En muchas aplicaciones de procesamiento de imágenes el objetivo es obtener una imagen donde sólo se mantengan los bordes más abruptos, mientras que el ruido y las pequeñas estructuras (artefactos) desaparezcan. La idea original de Perona y Malik [Perona, 1990] fue utilizar la difusión como un método de filtrado que preservaba los bordes. La difusividad térmica es sustituida por una función de difusividad que adapta la difusión a un gradiente local: difusión decreciente para gradiente creciente. De aquí que la ecuación de difusión pase a ser ahora

$$\frac{\partial I(x, y, t)}{\partial t} = \text{div}(D\nabla I(x, y, t)) \quad \text{Ecuación III-4}$$

- $I(x, y, t)$ es el valor de la señal en el instante t y en la posición (x, y)
- D es matriz de difusividad

Según el valor que tome D se tiene un tipo de difusión u otra:

- a) Difusión isótropa: si D es valor constante k .
- b) Difusión isótropa no lineal o anisótropa: si D es una función no lineal $g(\cdot)$.

3.2.2.3.1. Difusión isótropa

La utilidad de la difusión para aplicaciones de procesamiento de señales se encuentra en el caso de difusión anisótropa, sin embargo es interesante estudiar el caso isótropo como punto de partida para el posterior desarrollo de la difusión anisótropa.

La idea de esta aproximación es bastante simple. Se parte de una imagen original $I_0(x, y)$. A partir de ésta se derivan una serie de imágenes convolucionando a la imagen original con un filtro gaussiano $g(x, y; t)$ de varianza t ,

$$I(x, y, t) = I_0(x, y) * g(x, y; t) \quad \text{Ecuación III-5}$$

A mayor valor de t , denominado también parámetro escala-espacio, imágenes más borrosas.

Koenderink [Koenderink, 1984] dedujo que esta familia de imágenes dependientes del parámetro t , podía ser vista como la solución de la ecuación de difusión del calor

$$\frac{\partial I}{\partial t} = k \operatorname{div}(\nabla I) = k \left(\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \right) \quad \text{Ecuación III-6}$$

- k es la difusividad, es una constante
- Condición inicial: $I(x, y, 0) = I_0(x, y)$

Koenderink [Koenderink, 1984] estableció las propiedades más importantes que se le exigen al filtro que se utilice para suavizar la imagen:

1. *Causalidad*. Con esto se consigue que ningún detalle espurio se cree a ningún nivel de resolución, excepto aquellos ya existentes en la imagen original.
2. *Homogeneidad e isotropía*. El filtrado debe ser suave para que los cambios de intensidad se sucedan de forma continua conforme se realiza un desplazamiento ascendente en la pila de imágenes.

Por tanto, el efecto de la difusión isotropa es una transferencia de píxeles con valores altos a píxeles con valores bajos, es decir, tiene un efecto de suavizado, o lo que es lo mismo, de filtrado paso de baja.

Desventajas de la difusión isotropa:

La localización de los bordes en un determinado nivel de resolución varía con respecto a la verdadera posición en la imagen original. Al igual que pasaba con el *filtro gaussiano*.

3.2.2.3.2. Difusión anisótropa

A partir de esta desventaja de los filtros de difusión isotropa Perona y Malik [Perona, 1987] dedujeron las condiciones que cualquier método de generación de imágenes multiescalas debería cumplir:

- 1) *Causalidad*. No deben generarse detalles espurios nuevos en los distintos niveles de resolución.
- 2) *Localización inmediata*. En cada nivel de resolución las fronteras de las regiones deben ser abruptas.
- 3) *Suavizado por partes*. En todas las escalas, el suavizado del interior de las regiones debe ocurrir antes que el suavizado entre regiones.

Existe una forma fácil de modificar la difusión isotropa para que cumpla con las condiciones enunciadas arriba. En la difusión anisótropa el coeficiente de difusividad k , que hasta ahora se ha supuesto constante, es sustituido por una función $c(x, y, t)$. La idea es permitir la difusión a través de los bordes suaves y reducir la difusión para bordes abruptos.

En el caso ideal este procedimiento debería converger a una imagen en la que sólo se mantienen los bordes relevantes y donde el ruido y las pequeñas estructuras son eliminadas. La ecuación de difusión anisótropa viene dada por la siguiente ecuación diferencial en derivadas parciales [Perona, 1990]

$$\frac{\partial I}{\partial t} = \text{div}(c(x, y, t)\nabla I) = c(x, y, t)\Delta I + \nabla c \nabla I \quad \text{Ecuación III-7}$$

- Δ es el operador laplaciano
- si $c(x, y, t)$ es una constante, la ecuación se reduce a la de difusión isotrópica

Si en la escala t se conociesen las localizaciones de los bordes de las regiones, interesaría enfatizar el suavizado dentro de las regiones (coeficiente de difusividad a 1) frente al suavizado a través de los bordes (coeficiente de difusividad a 0). El problema se encuentra en que no se conoce a priori la situación de los bordes. Lo que se realiza es una estimación de la localización de los mismos para cada nivel de resolución a la que se denomina $\mathbf{E}(x, y, t)$, y que debe cumplir las siguientes propiedades:

- 1) $\mathbf{E}(x, y, t) = \mathbf{0}$ en el interior de cada región.
- 2) $\mathbf{E}(x, y, t) = K\mathbf{e}(x, y, t)$ en cada punto, donde \mathbf{e} es un vector unitario normal al borde en el punto (x, y) , y K es el contraste local del borde (diferencia entre las intensidades de la imagen en la vecindad del punto).

Si se dispone de una estimación de $\mathbf{E}(x, y, t)$, el coeficiente de difusividad $c(x, y, t)$ puede escogerse como una función de la magnitud de \mathbf{E} , es decir,

$$c = g(|\mathbf{E}|) \quad \text{Ecuación III-8}$$

Por tanto, para hallarlo hará falta definir primero cómo se calcula \mathbf{E} , y qué función no lineal $g(\cdot)$ se aplica:

- *Vector de bordes* $\mathbf{E}(x, y, t)$. El éxito del proceso de difusión está muy relacionado con la estimación que se haga de \mathbf{E} . Cuanto más precisa sea, más complicados serán los algoritmos. Perona y Malik [Perona, 1990] demuestran que la elección de \mathbf{E} como el gradiente de la función I (intensidad de la imagen) reporta excelentes resultados,
- *Función* $g(\cdot)$. Según las condiciones expuestas al principio de este apartado, debe ser una función no negativa, monótonamente decreciente y con $g(0) = 1$. De esta forma el proceso de difusión tiene lugar en el interior de las regiones y se dificulta en sus fronteras, donde la magnitud de \mathbf{E} es grande.

Perona y Malik [Perona, 1990] prueban con distintos tipos de funciones $g(\cdot)$, obteniendo resultados perceptivamente similares, y al final reducen las clases de funciones $g(\cdot)$, o lo que es lo mismo, los tipos de coeficientes de difusividad, a dos. Partiendo de

$$c(x, y, t) = g(|\nabla I(x, y, t)|) \quad \text{Ecuación III-9}$$

definen los dos coeficientes siguientes, según lo que se busque en la aplicación particular:

- 1) Cuando se quieren enfatizar los bordes de gran contraste frente a los de bajo contraste,

$$g(|\nabla I|) = \exp\left(-\left(\frac{|\nabla I|}{\gamma}\right)^2\right) \quad \text{Ecuación III-10}$$

- 2) Cuando se quieren enfatizar las regiones grandes frente a las pequeñas,

$$g(|\nabla I|) = \frac{1}{1 + \left(\frac{|\nabla I|}{\gamma}\right)^2} \quad \text{Ecuación III-11}$$

- γ representa un valor umbral

Difusión anisótropa para funciones que toman valores complejos

En esta apartado se explica el proceso de difusión anisótropa, pero aplicado a una función compleja. La motivación principal de este caso es el uso posterior de este tipo de difusión en la segmentación de imágenes en color, utilizando una notación conjunta del tono y la saturación mediante fasores. Esta notación fue introducida por [Lucchese, 2001].

Sea la función compleja

$$k(x, y) = \sigma(x, y)e^{j\theta(x, y)} \quad \text{Ecuación III-12}$$

- $\sigma(x, y)$ es el módulo en la posición (x, y)
- $\theta(x, y)$ es la fase en la posición (x, y)

La difusión anisótropa aplicada a $k(x, y, t)$ se define como

$$\frac{\partial}{\partial t} k(x, y, t) = \text{div}(c(x, y, t)\nabla k(x, y, t)) \quad \text{Ecuación III-13}$$

El gradiente viene dado por

$$\nabla k(x, y, t) = (\nabla \sigma(x, y, t) + j\sigma(x, y, t)\nabla \theta(x, y, t))\exp(j\theta(x, y, t)) \quad \text{Ecuación III-14}$$

y su magnitud

$$|\nabla(k(x, y, t))| = \sqrt{|\nabla(\sigma(x, y, t))|^2 + \sigma^2(x, y, t)|\nabla(\theta(x, y, t))|^2} \quad \text{Ecuación III-15}$$

Separando las partes real e imaginaria de la *Ecuación III-15*, se obtiene el siguiente para de ecuaciones acopladas

$$\begin{aligned} \frac{\partial}{\partial t} \sigma &= \text{div}(c\nabla \sigma) - c\sigma|\nabla \theta|^2 \\ \frac{\partial}{\partial t} \theta &= \text{div}(c\nabla \theta) + 2\frac{c}{\sigma}\nabla \sigma \cdot \nabla \theta \end{aligned} \quad \text{Ecuación III-16}$$

se ha omitido la dependencia espacial y temporal por la simplicidad en la presentación.

3.3. Métodos básicos de segmentación

A continuación se mostrará una visión general de los métodos más importantes en segmentación, tanto para imágenes monocromáticas, como sus extensiones a imágenes en color. Se presentan las tres aproximaciones básicas: *segmentación basada en el píxel*, *segmentación de regiones* y *detección de discontinuidades*, así como los diferentes métodos dentro de cada una de estas clasificaciones.

3.3.1. Segmentación basada en el píxel

La segmentación basada en el píxel considera solamente el valor de éste y no ningún tipo de relación con los demás píxeles de la imagen. Dentro de estas técnicas se encuentran la *umbralización* y los *algoritmos de búsqueda de agrupaciones*.

3.3.1.1. Umbralización

La umbralización es uno de los métodos más importantes en la segmentación de imágenes [González, 1996].

Son los métodos de segmentación más simples y usados [Plataniotis, 2000]. Se basan en el valor del píxel y no consideran el contexto espacial, es decir, deciden solamente en base a las características de color de los píxeles individuales.

3.3.1.1.1. Umbralización simple

Es la más sencilla de todas las técnicas de umbralización. Consiste en la partición del *histograma* de una imagen utilizando un umbral único T . El histograma de una imagen es una representación del número de píxeles que tiene un determinado valor frente a los valores posibles que toman. Propiamente normalizado, el histograma se puede definir como la función densidad de probabilidad de que un determinado valor ocurra.

La segmentación se completa entonces escaneando la imagen píxel a píxel y etiquetando cada píxel como perteneciente al objeto o al fondo, dependiendo de que el nivel de gris de ese píxel sea mayor o menor que el valor de T . El éxito de este método depende totalmente de cómo se pueda dividir el histograma [González, 1996].

Sea una función T

$$T = T(n, m, p_{n,m}, f_{n,m}) \quad \text{Ecuación III-17}$$

- $f_{n,m}$ es el valor del píxel en el punto (n,m)
- $p_{n,m}$ representa alguna propiedad local de ese punto

Una imagen normalizada $g_{n,m}$ se define como:

$$g_{n,m} = \begin{cases} 1 & \Rightarrow f_{n,m} > T \\ 0 & \Rightarrow f_{n,m} \leq T \end{cases} \quad \text{Ecuación III-18}$$

Por tanto, los píxeles marcados con 1 corresponden a objetos, mientras que los píxeles marcados con 0 corresponden al fondo.

Desventaja de la umbralización simple:

Este método es apropiado para imágenes que consistan en objetos brillantes sobre un fondo oscuro, y que por tanto, tengan un histograma bimodal. En este caso, se puede elegir el umbral para segmentar la imagen como aquel valor que corresponde con el valle del histograma.

3.3.1.1.2. Umbralización de mínimo error u óptima

Si se conocen las probabilidades a priori de los píxeles del objeto y del fondo y las distribuciones de las dos poblaciones, se puede elegir el umbral que exactamente minimiza el error.

Desventaja de la umbralización de mínimo error:

Se deben conocer las probabilidades a priori de los píxeles del objeto y del fondo y las distribuciones de las dos poblaciones.

3.3.1.1.3. Umbralización por el método de Otsu

Es una técnica de umbralización adaptativa para separar dos clases entre sí. Cada clase viene caracterizada por su centroide y por el área que forma la unión de los puntos asociados a dicho centroide. El algoritmo busca la maximización de la distancia entre las dos clases y la minimización de sus áreas [Weeks, 1996]. Este método no depende del modelado de las funciones de densidad de probabilidad [Petrou, 1999].

A continuación se muestra el planteamiento de esta técnica. Sea

- $(h(f_i))$ es el histograma de la imagen a umbralizar
- $m_1(K)$ es la media de los valores de aquellos píxeles cuyas intensidades estén por debajo de un umbral K (píxeles de la clase 1)

$$m_1(K) = \sum_{i=0}^K ih(f_i) \quad \text{Ecuación III-19}$$

- $m_2(K)$ es la media de los valores de aquellos píxeles cuyas intensidades estén por encima de un umbral K (píxeles de la clase 2)

$$m_2(K) = \sum_{i=K+1}^{f_{\max}} ih(f_i) \quad \text{Ecuación III-20}$$

- $P_1(K)$ es la probabilidad de ocurrencia de que un píxel pertenezca a la clase 1 en una imagen determinada

$$P_1(K) = \sum_{i=0}^K h(f_i) \quad \text{Ecuación III-21}$$

- $P_2(K)$ es la probabilidad de ocurrencia de que un píxel pertenezca a la clase 2 en una imagen determinada

$$P_2(K) = \sum_{i=K+1}^{f_{\max}} h(f_i) \quad \text{Ecuación III-22}$$

Se cumple que $P_1(K) + P_2(K) = 1$

- *Varianza entre clases* se obtiene a partir de

$$\sigma^2 = \frac{[m_1(K) \cdot P_2(K) - m_2(K) \cdot P_1(K)]}{P_1(K)P_2(K)} \quad \text{Ecuación III-23}$$

La maximización de la ecuación III-23 con respecto a K provee la mayor separación entre las dos clases y minimiza las áreas de cada una de ellas, obteniendo el umbral óptimo como

$$\hat{K} = \arg \left\{ \max_k \left\{ \frac{[m_1(K) \cdot P_2(K) - m_2(K) \cdot P_1(K)]}{P_1(K)P_2(K)} \right\} \right\} \quad \text{Ecuación III-24}$$

Desventajas de la umbralización por el método de Otsu:

- Supone que la media y varianza es suficiente para representar una función de densidad de probabilidad. Esto puede no ser cierto.
- El método no funciona cuando las dos distribuciones son muy distintas en tamaño entre sí.
- El método supone que el histograma de la imagen es bimodal, es decir, que la imagen contiene dos clases. Para más de dos clases tiene que ser modificado de forma que se definan múltiples umbrales.

3.3.1.1.4. Umbralización del histograma de homogeneidad (búsqueda de picos en el histograma de la imagen de un solo plano)

Existe un algoritmo de segmentación basado en la umbralización del histograma de homogeneidad. Este algoritmo utiliza un método de búsqueda de picos para obtener los picos del histograma. Este método será la base de la segmentación de algunos de los métodos presentados en este trabajo.

La homogeneidad está muy relacionada con la información local extraída de una imagen y refleja cómo de uniforme es una región. Se define como una composición de dos componentes: desviación estándar y discontinuidad de las intensidades I [Plataniotis, Venetsanopoulos, 2000]. La desviación estándar describe el contraste dentro de una región, mientras que la discontinuidad es una medida de los cambios abruptos en niveles de gris y se puede obtener aplicando detectores de bordes.

Para calcular el histograma de homogeneidad, se calcula el valor de homogeneidad para cada píxel. Posteriormente, para cada valor de intensidad, se suman los valores de homogeneidad de todos los píxeles con esta intensidad, tales que el valor de homogeneidad de ese píxel es igual o superior a 0.95 (son suficientemente homogéneos). Este valor se ha estimado experimentalmente.

Una vez calculado el histograma, se lleva a cabo un algoritmo de búsqueda de picos para seleccionar sólo los picos relevantes. Este algoritmo será el que utilicemos en este estudio. Los pasos a seguir son los siguientes:

1. Encontrar todos los puntos correspondientes a máximos locales en el histograma.
2. De los picos encontrados en la primera fase, volver a seleccionar los máximos locales. Los puntos resultantes son mucho más significativos que los del primer paso, para la determinación de los picos del histograma.
3. Umbralización: Se lleva a cabo en tres pasos. En primer lugar, si un pico es demasiado pequeño en relación con el pico más grande, se elimina. En segundo lugar, si dos picos están muy juntos, se elimina el que tenga menor valor. Por último, si el valle existente entre dos picos no es lo suficientemente significativo, también se elige sólo uno de los picos. Para examinar los valles, se calcula el valor medio de los valores entre los dos picos considerados. Si dicho valor medio es comparable a los valores de los picos, entonces se descarta el menor de ellos.

Una vez seleccionados los picos, las fronteras entre regiones serán los valores mínimos entre cada dos picos, es decir, los valles. Los valores de intensidad comprendidos entre cada dos valles formarán una región uniforme [Adame, 2002], [Acha, 2003].

En este estudio se hace uso del algoritmo de búsqueda de picos para poder encontrar el umbral con el cual segmentar una imagen de un solo plano de manera automática.

3.3.1.1.5. Umbrales basados en varias variables. Umbralización de imágenes en color

Anteriormente se ha tratado la umbralización de imágenes de una sola variable de intensidad. Puede plantearse el caso de tener que calcular umbrales basados en más de una variable [González, 1996].

El ejemplo más notable es el de imágenes en color, donde se utilizan los componentes rojo (R), verde (G) y azul (B) para formar una imagen de color compuesta. En este caso, cada píxel está caracterizado por tres valores, lo que permite la construcción de un histograma tridimensional [González, 1996].

El procedimiento más general para imágenes en color, es dividir el espacio tridimensional en celdas, representando cada una de ellas un color, y calcular el número de píxeles que toman un valor determinado. Dividiendo por el número total de píxeles de la imagen, se obtiene el histograma normalizado. El concepto de umbralización en este caso, se transforma en encontrar

agrupaciones de puntos en un espacio tridimensional [González, 1996]. El problema de búsqueda de agrupaciones es bastante más complejo que la simple umbralización.

Un método de segmentación que utilice los componentes tono y saturación de una señal de color es también particularmente atractivo, dado que implica agrupaciones de datos bidimensionales que resultan más fáciles de analizar que, por ejemplo, los tridimensionales que se necesitan para la segmentación en RGB [González, 1996]. En el plano *RGB* es difícil interpretar los resultados, sin embargo cuando se utiliza el plano *HS* (tono y saturación) los resultados obtenidos son mucho mejores, puesto que el sistema de representación según el tono y la saturación actúa como el sistema visual humano a la hora de percibir los colores. Otra desventaja del espacio de color *RGB* es que dispersa poco los valores de los colores y, por tanto, no facilita el establecimiento de umbrales para discriminar las diferentes regiones presentes [Russ, 1995].

3.3.1.2. Algoritmos de búsqueda de agrupaciones

Los métodos englobados en este apartado se centran en la propiedad de que generalmente los colores en una imagen tienden a formar *agrupaciones* en el histograma [Plataniotis, 2000].

Estas técnicas calculan primero el histograma a partir de los colores de los píxeles, y posteriormente se halla la forma de cada agrupación

Existen muchos algoritmos de búsqueda de agrupación de píxeles. Algunos de los más importantes son el algoritmo de *K-means* y el de *Fuzzy K-means*.

Estas técnicas no serán usadas en este estudio, por haber sido ya probadas en estudios anteriores [Acha, 2002].

3.3.1.3. Desventajas de la segmentación basada en el píxel

Sólo se consideran los valores de los píxeles. Sin embargo, la información espacial es muy importante para la segmentación de imágenes en las que existen píxeles del objeto del mismo color o nivel de gris que el fondo. Es decir, generalmente el problema de la segmentación necesita la información adicional de la localización del espacio. Por tanto, aunque los algoritmos son simples de implementar, en general, llevan a resultados pobres, sobretodo, en imágenes en color de escenas naturales

3.3.2. Segmentación orientada a regiones

Los métodos basados en regiones se centran en áreas del plano de la imagen en las que una propiedad de la imagen es homogénea [Heijden, 1994]. A diferencia de otras técnicas, tienen en cuenta tanto la distribución del color como la posición espacial de los píxeles. Son muy usados en la práctica debido a su simplicidad de implementación [González, 1996]. En muchos casos se aplican junto a otros tipos de algoritmos de segmentación, como por ejemplo, los basados en detección de discontinuidades.

Sin embargo, en este trabajo no se hará uso de este tipo de segmentación.

3.4. Segmentación basada en textura

Este tipo de segmentación se aplica a imágenes que contienen superficies de textura. “Textura” se define como “una superficie que consiste en un número grande, más o menos, de primitivas similares, que se encuentran dispuestas en un patrón de forma regular” [Heijden, 1994]. Regiones que no son uniformes en términos de los valores de grises o colores de sus píxeles, pero que se perciben como uniformes, se denominan *regiones de textura*. La textura puede verse como un descriptor que proporciona intuitivamente medidas de propiedades tales como suavizado, rugosidad y regularidad [González, 1996].

Para segmentar este tipo de imágenes, cada píxel no puede estar caracterizado solamente por su color, sino también por otro número o números que cuantifiquen la variación de los colores en un entorno alrededor del píxel.

Por tanto, cada píxel va a estar caracterizado por un vector de números, y cada componente del vector medirá alguna característica del píxel en dicha posición. Es decir, un píxel vendrá representado por un punto en un espacio multidimensional, donde la coordenada de cada eje es una medida de una *característica* [Petrou, 1999]. Los píxeles que pertenezcan a la misma región tendrán similares o idénticos valores en sus atributos y, por tanto, formarán una agrupación.

El problema se convierte en identificar agrupaciones de píxeles en un espacio multidimensional. Esencialmente el problema es el mismo que en la umbralización (apartado 3.3.1.1.) o en la búsqueda de agrupaciones (apartado 3.3.1.2.), salvo que ahora se trabaja con umbrales multidimensionales [Petrou, 1999].

3.5. Segmentación de imágenes en color. Espacios de color

Además de lo ya comentado de cada uno de los algoritmos de segmentación presentados sobre la ampliación a imágenes en color, en este apartado se comenta la segmentación en diferentes espacios de color.

3.5.1. Segmentación en el espacio de color HSI

Si se quiere segmentar una imagen en color y además de ello se quiere realizar el proceso en planos individuales, lo lógico sería pensar en un espacio de color HSI, puesto que el color es representado apropiadamente en el plano del tono. Generalmente, la saturación se utiliza como máscara para poder aislar más regiones de interés en la imagen del tono. La imagen de intensidad se usa menos frecuentemente en la segmentación porque no contiene información alguna sobre el color [González, Woods, 1996].

Utilizando este espacio de color los resultados obtenidos no son del todo satisfactorios. Utilizando el método de segmentación que se presenta en el siguiente apartado se pueden obtener resultados mucho mejores.

3.5.2. Segmentación en el espacio de color RGB

Aunque, como ya se ha dicho, trabajar en el espacio HSI es más intuitivo, cuando hablamos de segmentación, normalmente los mejores resultados se obtienen usando vectores RGB.

Supongamos que el objetivo es segmentar objetos de un determinado rango de color en una imagen RGB. A partir de un conjunto de puntos representativos de los colores de interés, se obtiene una estimación del “color medio” que se quiere segmentar.

- \mathbf{a} es el vector que representa el “color medio”.

El objetivo de la segmentación será clasificar cada píxel de la imagen dentro del rango especificado o no. Para poder realizar esta comparación, es necesario tener una medida de la similitud. Una de las formas de medida más simples es la distancia euclídea. Sea:

- \mathbf{z} representa un punto arbitrario en el espacio RGB.

Decimos que \mathbf{z} es similar a \mathbf{a} si la distancia entre ellos es menor que el umbral, D_θ . La distancia Euclídea entre \mathbf{z} y \mathbf{a} viene dada por

$$D(\mathbf{z}, \mathbf{a}) = \|\mathbf{z} - \mathbf{a}\| = \left[(\mathbf{z} - \mathbf{a})^T (\mathbf{z} - \mathbf{a}) \right]^{\frac{1}{2}} = \left[(z_R - a_R)^2 + (z_G - a_G)^2 + (z_B - a_B)^2 \right]^{\frac{1}{2}}$$

Ecuación III-25

- Los subíndices R, G y B indican las componentes del vector \mathbf{a} y \mathbf{z} .

Los puntos en los que se cumple $D(z, a) \leq D_0$ es una esfera sólida de radio D_0 . Los puntos dentro de la esfera o sobre ella satisfacen el criterio de color establecido, los puntos fuera de la esfera no. Codificando estos dos tipos de puntos en la imagen, por ejemplo blanco y negro, se obtiene una imagen segmentada binaria.

Una útil generalización de la *Ecuación III-25* es la medida de distancia según la ecuación

$$D(z, a) = \left[(z - a)^T C^{-1} (z - a) \right]^{\frac{1}{2}} \quad \text{Ecuación III-26}$$

- C es la matriz de covarianza de las muestras representativas del color que se desea segmentar

Otra opción sería trabajar con el cuadrado de la distancia, de esta forma se evitarían operaciones con raíces.

Un compromiso que se suele usar es la utilización de un cuadrado (“caja”) en lugar de la esfera. En esta aproximación, la “caja” está centrada en a . Dado un punto arbitrario, se segmentará determinando si dicho punto está fuera o dentro, o sobre la superficie, de la “caja” [González, Woods, 1996].

3.6. Resumen

En este capítulo se han repasado las principales herramientas utilizadas en la segmentación de imágenes. La segmentación trata de la reducción del contenido de información de una imagen para un más fácil procesamiento posterior por un sistema de visión por computador.

En general, todo algoritmo de segmentación empieza con un procesamiento inicial que prepara a la imagen para, posteriormente, segmentarla.

- a) Como herramienta básica del preprocesamiento se encuentra el *filtrado*.

Se han expuesto los principios que rigen al filtrado gaussiano, de mediana y de difusión. A este último se le ha prestado especial atención por su aplicación posterior en los algoritmos propuestos en esta tesis. El filtrado de difusión suaviza la imagen a la par que mantiene la información de los bordes, lo cual es muy deseable en aplicaciones de segmentación.

- b) La segunda herramienta de preprocesamiento es la *cuantificación*.

Esta herramienta permite disminuir el número de colores y, por tanto, el tiempo de computación de los algoritmos que se apliquen después. Sin embargo, en este trabajo no se hará uso de esta técnica.

El resto del capítulo trata de establecer una visión general de los métodos más importantes en segmentación, tanto para imágenes monocromáticas, como sus extensiones a imágenes en color.

Los métodos básicos de segmentación se clasifican en tres tipos:

- a) *Segmentación basada en el pixel*.

Considera solamente el valor de éste y no ningún tipo de relación con los demás píxeles de la imagen. Dentro de estas técnicas se encuentran la umbralización y los algoritmos de búsqueda de agrupaciones.

- b) *Segmentación orientada a regiones*

Estas técnicas intentan identificar un conjunto de píxeles coherentes espacialmente, que constituyen regiones uniformes en la imagen.

No serán utilizadas en este estudio.

c) *Detección de discontinuidades*

No han sido tratadas en este estudio.

Además de estas técnicas básicas también se realiza un estudio de otras técnicas de segmentación, como:

a) *Segmentación basada en textura*

Se considera la característica de textura. En este caso, cada píxel no puede estar caracterizado solamente por su color, sino también por otro número o números que cuantifiquen la variación de los colores en un entorno alrededor del píxel.

En último lugar se presenta un apartado dedicado a la segmentación de imágenes en diferentes espacios de color, tanto en el espacio HSI como en el RGB. Se muestran las ventajas que presenta este segundo respecto al primero, así como su formulación.

En el siguiente capítulo se resumen los trabajos más importantes realizados en el tratamiento digital de imágenes en color en los últimos años. Se estudiará la caracterización del grado de una quemadura y se hará el planteamiento del problema a resolver: Segmentación de quemaduras.
