C A P Í T U L O

SEGMENTACIÓN DE Imágenes Médicas

"Cuanto más se dividen los obstáculos son más fáciles de vencer". Concepción Arenal.

Profundizamos en este capítulo en la definición de Segmentación de imágenes digitales, y ponemos especial hincapié, como no podía ser de otra manera, en las imágenes en tres dimensiones.

Tras describir los principios generales del proceso de segmentación, llevamos a cabo una descripción exhaustiva de las distintas técnicas de segmentación tridimensional de imágenes médicas.

Por último, y como caso particularmente complejo, dedicamos una sección al estudio de la segmentación de tejidos blandos, por ser también uno de los objetivos del Proyecto VirSSPA.

3.1. Introducción

La Segmentación es la técnica de tratamiento digital de imágenes que permite extraer información de los objetos a partir de unas escenas dadas y plasmar esta información en un sistema estructurado, que en muchas ocasiones constará de una única estructura [9]. La Segmentación es una de las operaciones más usadas en el procesado digital de imágenes en tres dimensiones y es, casi con toda seguridad, la más difícil de todas ellas.

La Segmentación puede ser concebida como la consecución de dos tareas complementarias: el reconocimiento y la delineación. El **reconocimiento** consiste en la tarea de alto nivel de determinación a grandes rasgos de la ubicación del objeto en la escena estudiada. La **delineación** consiste en la tarea de bajo nivel de determinación precisa de la extensión espacial del objeto y su composición graduada. En cualquier caso, estos dos pasos suelen ser indisolubles, y hacemos esta distinción para entender conceptualmente las entrañas de la Segmentación digital. Pretendemos elaborar algoritmos que sean capaces de delinear los objetos de forma más precisa y eficiente que cualquier observador humano.

De acuerdo con lo dicho anteriormente, podemos dividir los algoritmos de **Reconocimiento** en dos grandes grupos: **automáticos** y **asistidos por el hombre**.

En los primeros, podemos optar por dos aproximaciones:

- Aproximación basada en conocimiento, que utiliza la inteligencia artificial para representar un conocimiento previo de los objetos de estudio en la región de interés. Usualmente se recurre a una delineación previa, con varias iteraciones, hasta poder comprobar las hipótesis realizadas.
- Aproximación basada en atlas, en la que un atlas de previa construcción representa la geometría y la relación topológica de los objetos en la región corporal de interés.

No obstante lo anterior, siempre será necesaria la supervisión de algún técnico que compruebe que no se han producido errores y que los parámetros han sido interpretados correctamente. Aún no existe el algoritmo que sea capaz de funcionar por sí solo sin ningún tipo de supervisión. En general el facultativo podrá encargarse de introducir la *semilla* inicial del algoritmo de acuerdo con los estudios realizados, o la indicación de una caja de trabajo que encierre al objeto que estemos considerando. También podrá aceptar o rechazar los objetos reconocidos en función de su existencia real o de que se trate de un fallo.

Es importante distinguir entre dos tipos de tareas de reconocimiento, el primero para identificar crecimientos patológicos como los que tienen lugar en lesiones en los senos y en los pulmones recogidas en imágenes de rayos X, y el segundo tipo para identificar los órganos anatómicos en las escenas. Las patologías son detectadas normalmente gracias al clásico patrón de reconocimiento en dos fases: extracción de características basada en la intensidad de la escena y clasificación de patrones. Desafortunadamente, este paradigma es inefectivo para la segunda tarea de reconocimiento de órganos en la escena. Será necesaria, por tanto, la participación del facultativo.

Por otra parte, y con respecto a la **Delineación**, las aproximaciones son estudiadas de forma más extensa que las de reconocimiento. De hecho, en muchas ocasiones los términos Segmentación y Delineación son uno mismo. Existen dos tipos de métodos de delineación: unos *basados en los bordes*, en los que el objeto de salida representa los bordes del objeto, y otros *basados en regiones*, en los que la estructura de salida representa la región ocupada por el objeto. En ambos tipos, la estructura de salida puede ser *dura (hard - crisp)* o *borrosa (fuzzy)*. En el primer caso la estructura es un conjunto de elementos para los que el grado de pertenencia puede ser 0 ó 1. En cambio, en el caso borroso, la estructura es un conjunto de valores para los que el grado de pertenencia va de 0 a 1 dependiendo del elemento.

En virtud de la clasificación realizada anteriormente, podemos establecer siete tipos de Segmentación, de los que haremos una revisión completa a continuación [9].

3.2. Clasificación de técnicas de segmentación 3D

3.2.1. Métodos de decisión dura, basados en bordes y automáticos

Existen principalmente dos aproximaciones para la creación de representaciones de decisión dura basadas en bordes de los objetos procedentes de escenas de la región corporal que estamos analizando. En el primer grupo, se asume que una superficie del objeto puede quedar determinada por un umbral fijo de intensidad de la escena. Esto es, la estructura de salida que representa el objeto es la superficie de la escena origen cuyo valor de intensidad es el escogido. A este grupo de métodos los denominamos de **isosuperficie**. En el segundo grupo, el gradiente o tasa de variación de la intensidad de la escena es utilizado para localizar y definir los bordes. Nos referimos a éstos como métodos basados en **gradiente**.

3.2.1.1. Métodos de isosuperficie

A su vez, y agrupando, podemos distinguir entre dos grupos de aproximaciones. El primero de ellos formado por los **métodos de superficie digital**, que obtienen una representación digital de la superficie objeto. Aquí el término *digital* denota que que la representación está formada por elementos básicos, que adoptan una de entre un conjunto limitado de formas y orientaciones. Para el caso 3D, por ejemplo, los elementos básicos pueden ser vóxeles o las caras de los vóxel. Si el espacio entre láminas es igual entre todas las láminas de la escena dada, entonces todos los vóxeles de la escena son de las mismas dimensiones. Hay seis tipos de caras de vóxel basadas en su orientación. Son todas rectangulares, y su tamaño depende de la relación entre el tamaño del píxel y el espacio entre láminas de la escena.

El segundo grupo obtiene la superficie objeto como un conjunto de elementos triangulares o poligonales. Éstos no son usualmente digitales en el sentido descrito anteriormente. Los denominamos **métodos de superficie poligonal**.

Superficies digitales

Comencemos pues con la definición matemática formal de una superficie digital como el borde existente entre un objeto digital - un conjuntos de vóxeles conexos cuyas intensidades son mayores o iguales que el umbral de intensidad especificado T - y un co-objeto digital - un conjunto de vóxeles cuya intensidad está por debajo del umbral T. En una escena $\mathcal{V} = (V,g)$, la superficie común entre dos vóxeles v_1 y v_2 (que difieren exactamente en una coordenada) es un elemento del borde si $g(v_1) \geq T$ y $g(v_2) < T$. La orientación de este elemento viene determinada por el vector unitario en la dirección que va desde el centro de v_1 hasta el centro de v_2 . Las seis posibles orientaciones de los elementos del borde se corresponden con los vectores unitarios de las seis direcciones principales de los ejes coordenados.

Т

$$\bigcup_{i} v_i \mid g(v_i) \ge T \tag{3.1}$$

$$\bigcup_{i} v_i \mid g(v_i) < T \tag{3.2}$$

Una vez definido lo anterior, hemos de añadir que las superficies han de satisfacer las siguientes tres propiedades:

- Clausura: las superficie divide la escena en dos conjuntos, uno interior y otro exterior, de modo que ningún camino que parta de un vóxel en el interior y llegue a un vóxel en el exterior se encuentre con la superficie.
- Orientabilidad: la propia superficie tiene un interior y un exterior. La orientación del elemento borde indica la orientación de la superficie localmente.
- Conectividad: ningún elemento borde en la superficie puede ser alcanzado desde cualquier otro elemento borde que vaya desde un elemento hasta el adyacente sin dejar la superficie.

Los algoritmos basados en isosuperficie digital cumplen con estas propiedades. A continuación presentamos uno de los algoritmos más eficientes para encontrar isosuperficies cerradas, orientadas y conectadas.

- 1. Encontrar una cara inicial y situarla en una cola Q.
- 2. Si Q está vacía, paramos. Si no, extraemos una cara F de Q, salida F, y encontramos su única cara adyacente F' en cada dirección que sea asignada a F. La(s) dirección(es) asignada(s) a una cara y cómo encontrarlas se ilustran en la figura 3.1.
- 3. Para cada F' localizada en 2, si no fue encontrada previamente, ha de ser puesta en la cola Q. Ir al paso 2.



Figura 3.1: Forma más eficiente para encontrar isosuperficies cerradas, orientadas y conectadas [9].

Nótese que, en cualquier superficie bordeada, como en el modelo de la figura 3.1, sólo aquellas caras que son ortogonales al eje z del sistema cartesiano de la escena tienen dos direcciones asociadas con ellas, y el resto tienen únicamente una. Esto implica que la mayor parte del tiempo el algoritmo va desde una cara hasta una única cara adyacente hasta el momento en que se bifurca hacia dos caras adyacentes hasta asegurarse de que todas las caras pertenecientes a la superficie cerrada y conectada se han encontrado. Si el nivel de discretización (que viene dado por el tamaño de las caras de salida) deseado para la superficie de salida es diferente del de la escena de entrada entonces es sencillo incorporarla al paso 2, donde las intensidades de v_3 y v_4 se estiman a partir de una interpolación local.

Superficies poligonales

Entre estos métodos, destacan aquéllos que usan como elemento superficial básico un triángulo. Este hecho se deduce de la disponibilidad de estructuras triangulares en software estandarizado como Open GL, así como en aceleradores de hardware y motores de renderizado.

Los métodos más utilizados son los denominados *cubos desplazantes* y sus derivados. La idea que subyace bajo este planteamiento viene ejemplificada en la figura 3.2. Sea $\mathcal{V} = (V, g)$ una escena dada y sea T un umbral especificado. Consideremos un bloque de $2 \times 2 \times 2$ vóxeles donde v_1, v_2, \ldots, v_8 representan los ocho centros de los vóxeles del bloque. Cada uno de estos vóxeles puede pertenecer o no a la superficie borde que queremos construir, dependiendo de si $g(v_i) \geq T$ o $g(v_i) < T$. En la configuración de la figura

se cumple que $g(v_1) \ge T$ y $g(v_i) < T$ para i = 2, ..., 8. En este caso sólo v_1 pertenece a la superficie. Construimos por tanto la superficie localmente con un triángulo como el mostrado en la figura. En este caso, la localización de P_1 a lo largo de la línea que una v_1 y v_2 queda determinada por interpolación lineal en el lugar en el que la intensidad es exactamente igual a T. P_2 y P_3 quedan determinados análogamente.



Figura 3.2: En la figura podemos observar cómo se constituyen los llamados *marching cubes* [9].

Queremos resaltar que el método así constituido no garantiza el cumplimiento de las tres propiedades explicadas en este capítulo, y esta garantía aún no existe para ningún método. Existen nuevas aproximaciones que intentan garantizar el cumplimiento de las tres propiedades (no se garantiza la conectividad).



Figura 3.3: Ejemplo de Marching Cubes.

3.2.1.2. Métodos basados en el gradiente

La premisa básica de estos métodos consiste en que, en la vecindad del borde del objeto, la magnitud del gradiente de la intensidad de la escena es mucho mayor que en puntos alejados del borde. Así pues, el conjunto de elementos pertenecientes al borde - píxeles, vóxeles, lados de píxel y caras de vóxel - en los que la magnitud del gradiente de la intensidad sea suficientemente alta, representará el borde del objeto que estamos extrayendo. Desafortunadamente, en la práctica, este conjunto contiene elementos que que no pertenecen a la estructura y/o faltan elementos que verdaderamente sí pertenecen. Para superar este problema, existen diversas estrategias de optimización.

Una posible solución es asignar un coste a cada elemento del borde y encontrar un borde cuyo coste total es el menor entre todos los posibles. Este método exige un elevado poder computacional. Por eso, se suele optar por realizar suposiciones derivadas de la naturaleza de los bordes y de la observación. Otra solución es comenzar por un elemento del borde cualquiera e ir añadiendo, de todos los vecinos posibles, aquél que esté más cualificado cada vez.

3.2.2. Métodos borrosos, basados en bordes y automáticos

En este caso, los conceptos de clausura, orientabilidad y conectividad dejan de tener valía. Hablamos de **medida de pertenencia al borde** para cada vóxel de la escena sin preocuparnos de estos tres conceptos.

Una forma muy simple de atender a estos métodos es definir una función (usualmente lineal) de pertenencia al borde, función de la magnitud gradiente $|\vec{\nabla}g(v)|$ y especificada dentro de un intervalo de $|\vec{\nabla}g(v)|$. Por encima y por debajo de este intervalo la pertenencia al borde puede quedar determinada por la asignación de un valor alto y bajo, respectivamente.

A veces la intensidad de la escena g(v) y el módulo del gradiente $|\nabla g(v)|$ se usan para determinar la pertenencia al borde. Por ejemplo, especificando dos rectángulos anidados en un histograma bidimensional de la intensidad y el gradiente. La pertenencia al borde es 1 dentro del rectángulo interior y 0 fuera del rectángulo exterior, y varía de forma bilineal entre ambos.

3.2.3. Métodos de decisión dura, basados en bordes y asistidos

Es evidente que el más simple de todos estos métodos es aquél que delimita las líneas externas lámina a lámina de los objetos de los que queremos determinar el borde. Cuando las regiones del objeto tienen múltiples contornos y/o agujeros en las láminas, esta idea asume un reto que ha de implementar, especialmente cuando queremos obtener una representación binaria de la escena con la estructura segmentada.

3.2.3.1. Active Contours

Con este nombre se agrupan un conjuntos de métodos asistidos más sofisticados, cuya principal intención es minimizar el grado de asistencia que es necesario prestar al algoritmo. También se los conoce como bordes activos, serpientes, superficies activas y métodos de bordes deformables.

Aunque existen métodos para el caso tridimensional, vamos a intentar comprenderlo mediante el caso 2D. En primer lugar, se especifica un contorno inicial que se encuentra en la vecindad de la superficie de interés. Este borde se va deformando subsecuentemente hasta quedar confinando los bordes de la superficie de la imagen. Se establece un sistema mecánico donde los bordes de la imagen parecen tener cierta rigidez, y se ejercen las fuerzas sobre el contorno en puntos característicos de la imagen, normalmente bordes. La forma del contorno final se estima minimizando una funcional de energía, formada por una componente de energía externa y una componente de energía interna. La energía interna impone suavidad y restricciones de continuidad en el contorno. La energía externa atrae el contorno hasta los bordes de la imagen. El contorno se representa a través de una curva paramétrica v(s) = [x(s), y(s)] donde *s* es la longitud del arco. La funcional de energía se define como

$$S(u,v) = \left[x(u,v), y(u,v), z(u,v)\right]$$
(3.3)

$$E = \int_0^1 \int_0^1 \left\{ E_{int} \left[S(u, v) \right] + E_{ext} \left[S(u, v) \right] \right\} du \, dv \tag{3.4}$$

$$E = \int_0^1 \left\{ E_{int}[v(s)] + E_{ext}[v(s)] \right\} ds$$
 (3.5)

donde E_{int} y E_{ext} son la energía interna y externa respectivamente. *E* ha de minimizarse para determinar la curva. La diferencia entre los distintos métodos viene dada por:

- 1. Cómo se especifica el contorno inicial.
- 2. Cómo se definen los bordes
- 3. Cómo se definen la energía interna y externa.
- 4. Cómo se halla el mínimo de la funcional *E*.

Normalmente, E_{int} contiene términos que regulan la elasticidad y rigidez del contorno y E_{ext} contiene términos que regulan las fuerzas y restricciones especificadas por el usuario.



Figura 3.4: Ejemplo de aplicación de Active Contours en tres dimensiones.

La precisión de estos métodos no ha sido evaluada con cuidado, aunque se utilizan en multitud de aplicaciones. Dedicaremos un capítulo especial a este método de Segmentación, explicado el porqué de su elección.

3.2.3.2. Live wire/lane

En las aproximaciones derivadas de la segmentación Active Contours no podemos garantizar que la superficie correspondiente al mínimo en la funcional de energía sea la correcta. Por tanto será necesaria una examinación manual y una corrección para un estudio básico. Para evitar esto, se ha desarrollado recientemente un paradigma alternativo, que ha dado lugar a los llamados **métodos live wire/lane**. En ellos, un operador ofrece un reconocimiento continuo al algoritmo de delineación automático. Además, estos métodos difieren de Active Contours en la optimización del modelo, como explicamos a continuación.



Figura 3.5: Modelo de grafo usado en el método Live Wire. Los vértices de los píxeles son los nodos del grafo, y cada lado de los píxeles queda determinado por dos caminos diferentes [9].

Queremos conseguir la superficie borde de una vez sin necesidad de hacer correcciones posteriores. Para conseguirlo, el usuario debe prestar un férreo control al proceso de segmentación mientras éste está en vías de ejecución. El usuario selecciona inicialmente un punto de la superficie de interés sobre el display de una lámina. En ese momento, un segmento live wire - una curva digital óptima - desde el punto seleccionado hasta otro punto subsecuente indicado por el cursor se computa y se muestra en tiempo real.Si el cursor está cerca de la superficie y no muy lejos del punto inicial, entonces el segmento live wire roza el borde. Si el segmento delinea la frontera exactamente, el cursor queda fijado en ese punto, que se convierte en un nuevo punto de partida. En la práctica, hemos de especificar de dos a cinco puntos iniciales.

3.2.4. Métodos de decisión dura, basados en regiones y automáticos

En estos métodos, podemos considerar como salida una escena binaria $\mathcal{V}_o = (V_o, g_o)$ en la que los vóxeles con intensidad 1 representan la región ocupada por la estructura. La entrada es un conjunto de escenas $\mathcal{V}_{ij} = (V_{ij}, g_{ij})$ para $i \leq j \leq s$ correspondientes a la misma región corporal. Podemos identificar dos clases de métodos: **umbralización** y **clustering** o agrupación.

3.2.4.1. Umbralización

En estos métodos, se escoge un intervalo umbral $[t_{j1}, t_{j2}]$ para cada escena de entrada de modo que la estructura pueda ser especificada completamente por los vóxeles cuyas intensidades recaigan en este intervalo. La escena binaria de salida se define de la siguiente forma:

$$V_o = V_{i1} = V_{i2} = \dots = V_{is} \tag{3.6}$$

y para cualquier $v \in V_o$,

$$g_o(v) = \begin{cases} 1 & \text{si } g_{ij}(v) \in [t_{j1}, t_{j2}] \forall 1 \le j \le s \\ 0 & \text{en otro caso} \end{cases}$$
(3.7)

La umbralización es quizás el método de segmentación más utilizado por su simplicidad y se eficiencia. No obstante, destacamos que los resultados de una umbralización rara vez son perfectos, aunque sí puede resultar muy útil como una primera aproximación.



Figura 3.6: Ejemplo de Umbralización 3D para segmentar tejido óseo.

3.2.4.2. Clustering

La premisa básica de las técnicas de **clustering** consiste en la separación evidente de las distintas estructuras en racimos de puntos pertenecientes a un espacio característico de valores de las propiedades. Estas propiedades pueden ser la intensidad de la escena o propiedades derivadas de la misma, como el gradiente de intensidad de escena. Para identificar y definir estructuras, se identifican racimos en el espacio característico, y por tanto la segmentación pasa por ser la identificación de estos racimos. Los diversos métodos difieren entre sí en las aproximaciones adoptadas para esta identificación y definición de racimos.

Una aproximación más sofisticada, que por otro lado quizás sea el método más recurrente en aplicaciones médicas, es el **Método kNN** (the k nearest neighbour). Consideremos el caso de un objeto y un fondo y *m* escenas de entrada que representan distintos valores de propiedades en la misma región corporal. El método se divide en una fase de *entrenamiento* y en otra de *clasificación*.

- Identificar dos conjuntos de vóxeles, X_o en la región del objeto y X_b en la región del fondo. Esto puede ser realizado, por ejemplo, por un experto que dibuje las regiones en una pantalla de las láminas de las escenas obtenidas para una misma región corporal y el protocolo que se utiliza para adquirir las escenas dadas. Este paso de entrenamiento debe ser ejecutado sólo una vez al comienzo del proceso y no para cada estudio del paciente. Es necesario escoger un valor para el parámetro k.
- 2. Para cada vóxel v en la entrada $\mathcal{V}_{ij} = (V_{ij}, g_{ij}), 1 \leq j \leq m$, encontrar su localización $P(v) = (g_{i1}(v), g_{i2}(v), \dots, g_{im}(v))$, en el espacio característico.
- 3. Encontrar k vóxeles v' de la unión $X_o \bigcup X_b$ tales que P(v') sean los puntos más próximos en el espacio característico a P(v).
- Si la mayoría de estos vóxeles pertenece a X_o, clasificamos v como perteneciente al objeto. En otro caso, consideramos v perteneciente al fondo.

En todos los métodos clustering hay parámetros que han de ser previamente especificados. En el método kNN estos parámetros son k, X_o , X_b . Si no existe cierta asistencia, como ya indicamos en el apartado 1, no se puede garantizar la eficacia del método.



Figura 3.7: Ejemplo de Clustering en los distintos huesos componentes de la mano.

3.2.5. Métodos borrosos, basados en regiones y automáticos

Estos métodos pueden ser entendidos como la consecución de una escena de grises $\mathcal{V}_o = (V_o, g_o)$ en la que la intensidad de cada vóxel $g_o(v)$ representa el grado de pertenencia a objeto de v. La entrada es un conjunto de escenas de grises $\mathcal{V}_{ij} = (V_{ij}, g_{ij})$ para $1 \le j \le s$, para la misma región corporal.

El más simple entre estos métodos es la **umbralización borrosa**, de la que ilustramos un ejemplo en la figura 3.8 y que es una generalización de la umbralización de decisión dura. Comúnmente se usa con una única escena de entrada, $\mathcal{V}_i = (V_i, g_i)$. En este caso se requiere la especificación de cuatro umbrales de intensidad. La escena de salida se define como sigue:

$$V_o = V_i \tag{3.8}$$

para cualquier $v \in V_o$

$$g_o(v) = \begin{cases} 0 & \text{si } g_i(v) > t_4 \ o \ g_i(v) < t_1 \\ 1 & \text{si } t_2 \le g_i(v) \le t_3 \\ \frac{g_i(v) - t_1}{t_2 - t_1} & \text{si } t_1 \le g_i(v) \le t_2 \\ \frac{t_4 - g_i(v)}{t_4 - t_3} & \text{si } t_3 \le g_i(v) \le t_4 \end{cases}$$
(3.9)

Esto es, si la intensidad está entre los umbrales t_2 y t_3 tiene un 100 % de pertenencia al objeto. Si está fuera del intervalo $[t_1, t_4]$, no pertenece al objeto.



Para valores intermedios, la pertenencia al objeto toma valores entre 0 y 1.

Figura 3.8: Ejemplo de Umbralización Borrosa.

3.2.6. Métodos de decisión dura, basados en regiones y asistidos

La más simple de estas técnicas consiste en el perfilado a mano de las regiones de interés usando el driver de un ratón como pincel de selección. En aplicaciones médicas, muchas veces la traza del borde consume menos tiempo que la determinación de regiones.

En contraposición con estos métodos totalmente manuales, existen algunos en los que el reconocimiento es manual pero la delineación es automática. El **crecimiento de regiones** es un método muy popular en este grupo. En general, este método puede resumirse en los siguientes pasos:

- 1. Se establece un conjunto de criterios para la inclusión de vóxeles en la estructura. Algunos ejemplos son:
 - *a*) La intensidad de la escena de un vóxel ha de estar en el intervalo $[t_1, t_2]$.
 - *b*) La intensidad media de los vóxeles incluidos en la región de crecimiento en cualquier momento durante el proceso de crecimiento debe estar en el intervalo $[t_3, t_4]$.
 - *c*) La varianza de la intensidad de los vóxeles incluidos en la región de crecimiento en cualquier instante durante el proceso de crecimiento debe estar en el intervalo $[t_5, t_6]$.

- El usuario especifica uno o más vóxel semilla en el interior de la región del objeto que queremos segmentar usando el puntero del ratón en un display de láminas. Se colocan en una cola y son especialmente señalados.
- 3. Si la cola está vacía, se acaba el algoritmo. Si no, se borra el el vóxel *v* de la cola, y sacamos *v*.
- 4. Se examinan los vecinos (usualmente los 6, 18 o 26 más cercanos) de v para la inclusión. Aquéllos que satisfagan el criterio y no hayan sido previamente marcados y encolados son añadidos y marcados especialmente. Volver al paso 3.

Queremos hacer notar que, si usamos sólo el criterio a, y si t_1 y t_2 permanecen fijos durante el proceso de crecimiento, entonces el método anterior obtiene el mayor número de vóxeles conectados a la semilla inicial, satisfaciendo el intervalo dado. Para cualquier combinación de a hasta c, y si los valores de t_1, \ldots, t_6 no son fijos, no podemos garantizar que el conjunto de vóxeles $O(v_1)$ obtenido con la semilla v_1 es el mismo que el conjunto $O(v_2)$ obtenido a partir de la semilla $v_2 \neq v_1$. Esta falta de robustez es un problema en la mayoría de algoritmos de crecimiento de regiones.

3.2.7. Métodos borrosos, basados en regiones y asistidos

Incluimos en esta sección los métodos borrosos basados en regiones que, aún teniendo alguna fase automática, precisan de asistencia humana para un correcto diagnóstico. Por ejemplo, el método KNN cuando precisa de entrenamiento para cada estudio.

Un método desarrollado recientemente, que en el diseño incluye asistencia humana, es la **técnica de conectividad borrosa**. En este método, el reconocimiento es manual e implica señalar la ubicación de un objeto en el display de láminas. La delineación, en cambio, es automática. Este método ha sido utilizado en varias aplicaciones, tales como cuantificación de lesiones a partir de resonancias magnéticas, Angiografía MR y planificación para cirugía maxilofacial.

Pensamos en los vóxeles cercanos a la escena $\mathcal{V} = (V, g)$ que tienen una relación de afinidad . La fuerza de esta relación entre cualesquiera dos vóxeles v y u pertenecientes a V es una función de la distancia entre los dos vóxeles y de la intensidad de la escena en ellos. Por ejemplo,

$$\mu(v,u) = \frac{1 - |g(v) - g(u)|}{||v - u||}$$
(3.10)

donde ||v - u|| denota la distancia entre v y u, además de haber asumido que el rango de intensidades de la escena se encuentra en el intervalo [0, 1].

La función de afinidad, por otro lado, expresa el grado en el que los vóxeles se relacionan entre sí. Los vóxeles que están alejados muestran una afinidad imsignificante. En este sentido, definimos la función de afinidad de la siguiente manera:

$$\mu(v,u) = w_1 g_1 \Big(\vec{g}(v), \vec{g}(u) \Big) + w_2 \Big(1 - g_2(\vec{g}(v), \vec{g}(u)) \Big)$$
(3.11)

$$g_1\left(\vec{g}(v), \vec{g}(u)\right) = \frac{1}{(2\pi)^{m/2} |S_1|^{1/2}} e^{-\frac{1}{2} \left[\frac{1}{2} (\vec{g}(v) + \vec{g}(u)) - \vec{m}_1\right]^T S_1^{-1} \left[\frac{1}{2} (\vec{g}(v) + \vec{g}(u)) - \vec{m}_1\right]}$$
(3.12)

$$g_2\left(\vec{g}(v), \vec{g}(u)\right) = \frac{1}{(2\pi)^{m/2} |S_2|^{1/2}} e^{-\frac{1}{2} \left[\frac{1}{2} |\vec{g}(v) - \vec{g}(u)| - \vec{m}_2\right]^T S_2^{-1} \left[\frac{1}{2} |\vec{g}(v) - \vec{g}(u)| - \vec{m}_2\right]}$$
(3.13)

En estas expresiones, g_1 y g_2 son funciones gaussianas de varias variables, $\vec{g}(v)$ y $\vec{g}(u)$ son vectores columna con m componentes, $\vec{m_1}$ y $\vec{m_2}$ son vectores media de m componentes, S_1 y S_2 son matrices de covarianza de orden m, y $|\vec{g}(v) - \vec{g}(u)|$ denota la distancia absoluta entre las componentes centrales de $\vec{g}(v)$ y $\vec{g}(u)$. Finalmente w_1 y w_2 son pesos positivos que cumplen la condición $w_1 + w_2 = 1$.

De esta forma estamos aplicando de nuevo el concepto de *conectividad borrosa, K,* de la siguiente manera. Para cada posible par (v, u) de vóxeles en el dominio de la escena, una fuerza de conectividad $\mu(v, u)$ es asignada siguiendo el siguiente criterio. Hay diversos caminos posibles entre v y u, entendiendo camino como la secuencia de vóxeles que comienza en v y acaba en u y todos los vóxeles cercanos con cierto grado de adyacencia. La **fuerza** del camino es la menor de las afinidades asociadas a un par de vóxeles sucesivos a lo largo del camino. La fuerza de conectividad entre v y u es simplemente la mayor de las fuerzas asociadas en todos los caminos existentes entre v y u. De este modo, creamos un objeto borroso como una alianza de vóxeles, de modo que éstos cumplan que la fuerza de conectividad entre cualesquiera dos vóxeles en la alianza sea mayor que un umbral establecido y mayor que la fuerza de conectividad entre un vóxel interior al objeto y otro vóxel exterior.

Es evidente que aplicar la definición de afinidad dada en las ecuaciones anteriores es ciertamente complejo computacionalmente hablando, por tanto tendremos que realizar las simplificaciones y consideraciones necesarias, que se estudian profusamente en la bibliografía adjuntada.

3.3. Segmentación de Tejidos Blandos

3.3.1. Introducción

La segmentación de tejidos blandos requiere, en ocasiones, de técnicas específicas que consigan distinguir correctamente la región corporal en que se encuentran ubicados. A continuación presentamos las distintas técnicas y metodologías que pueden ayudar a la herramienta de planificación quirúrgica VirSSPA a segmentar correctamente órganos, tejidos blandos y tumores que, sin la utilización de las mismas, difícilmente podrían ser determinados.

En primer lugar, presentamos el problema y marcamos las dificultades con las que un algoritmo de segmentación convencional, basado en niveles de gris de la imagen, podría encontrarse. En segundo lugar, presentamos distintas técnicas de segmentación que podrían resolver el problema presentado anteriormente, haciendo una breve descripción de los mismos. En último lugar, obtenemos una serie de conclusiones que pueden sentar las bases para la elección del algoritmo definitivo que habrá de implementarse en la herramienta VirSSPA.

3.3.2. Presentación del problema

Todos los tejidos corporales no óseos, como los músculos, la grasa, el tejido fibroso, los vasos sanguíneos o cualquier otro tejido conjuntivo del cuerpo se denominan tejidos blandos. Desde un punto de vista técnico, los niveles de gris de una imagen TAC que contenga distintos tipos de estos tejidos no son lo suficientemente divergentes como para que el proceso de segmentación de los mismos pueda utilizar esta característica, el nivel de intensidad de la imagen, como único parámetro válido.

El paso de preprocesado incluido en el algoritmo, aun aumentando notablemente el rango dinámico de las regiones de interés dentro de la imagen, y por tanto optimizando su resolución y la divergencia entre los diferentes niveles de gris de los diferentes tejidos, no consigue que el tratamiento de tejidos blandos pueda tener en cuenta únicamente el nivel de gris de los vóxeles. Será necesario atender a parámetros estadísticos y, en general, será necesaria también la intervención del facultativo, que marcará, de nuevo, una serie de parámetros iniciales que nos permitan determinar la región definitiva.

3.3.3. Técnicas destacadas en Segmentación de Tejidos Blandos

De entre los trabajos seleccionados para la elaboración de este informe sobre el Estado del Arte de la Segmentación de tejidos blandos, dos de ellos destacan por la generalidad de las metodologías empleadas y por la extensa y precisa descripción de los mismos. Es por esto que hemos considerado interesante desarrollarlos como dos posibles alternativas para la elaboración de un algoritmo de segmentación de tejidos blandos para VirSSPA. Estos trabajos son el realizado por Heimann et al. [10] y el realizado por Dehmeshki et al. [11], el primero de ellos basado en técnicas *Active Contours* y el segundo de ellos en Crecimiento de Regiones y mapeado de conectividad borrosa. Dedicaremos sendas secciones para describirlos más extensamente.

Por otro lado, resulta especialmente interesante el trabajo de Feng et al. [12], que explica un modelo estadístico para la segmentación del tejido vascular, que podría ser notablemente interesante para patologías de origen cardiovascular. Además, hemos prestado atención al artículo presentado por Ritthipravat et al. [13], en el que utilizan un algoritmo de segmentación basado en modelo estadístico para carcinomas nasofaríngeos en imágenes TAC.

Además, el trabajo de Maeda et al. [14] muestra también la importancia de la utilización de las técnicas basadas en *Active Contours* para la segmentación de tejidos blandos en imágenes TAC. En el trabajo realizado por Wiegert et. al [15] lo más interesante es la construcción del modelo tridimensional, puesto que la técnica de adquisición de imágenes requiere un tratamiento sensiblemente diferente al necesario para VirSSPA.

Por último, destacamos la incorporación de imágenes PET, junto a TAC, propuesta por Lee et al. [16] así como el riguroso estudio realizado por Palumbo et al. [17] en el caso de la fibrosis quística pulmonar.

3.3.4. Modelo estadístico deformable

Heimann et al. [10] proponen un modelo estadístico deformable como solución al problema de la segmentación tridimensional de tejidos blandos. En primer lugar, exponen los parámetros que conforman el modelo estadístico, y posteriormente se explica cómo este modelo se adapta a la región de interés, a través de la definición de fuerzas externas e internas, respondiendo al modelo clásico de la segmentación *Active Contours*.

Modelo estadístico

A su vez, el modelo estadístico se construye con un conjunto de imágenes de entrenamiento y está compuesto de un modelo geométrico que describe la forma y un modelo de apariencia local que describe el contorno.

El **modelo geométrico** queda representado por una distribución de puntos, esto es, una densa colección de puntos singulares en la superficie del objeto. Cada forma de entrenamiento t queda determinada por un vector único x_t de coordenadas singulares concatenadas. La media de todos los vectores produce la forma media $\overline{\mathbf{x}}$ y un análisis en componentes principales de la matriz de covarianza produce los modos principales de variación p_m y sus respectivas varianzas λ_m . Haciendo uso de un número finito c de parámetros y_m , todas las formas válidas \mathbf{x} pueden ser aproximadas mediante la expresión

$$\mathbf{x} \cong \overline{\mathbf{x}} + \sum_{m=1}^{c} y_m \mathbf{p}_m \tag{3.14}$$

Un prerrequisito para la construcción del modelo de forma es que todos los puntos singulares estén situados en el lugar correspondiente en todas las muestras de entrenamiento. Para tal objetivo habrá que decidir posteriormente cómo se lleva a cabo el registro de las imágenes de modo que se obtenga una distribución homogénea de puntos.

Para detectar la forma anteriormente modelada en una nueva imagen, como ya hemos comentado, se emplea un **modelo de apariencia local** del nivel de gris del contorno. Un método común para efectuar esta operación es muestrear perfiles g_{it} perpendiculares a la superficie en cada punto singular *i* y en todas las imágenes de entrenamiento *t*. De nuevo podemos usar análisis en componentes principales para extraer los perfiles medios \overline{g}_i y sus varianzas.

Debido a la no linealidad inherente en las imágenes médicas, puede que el método descrito anteriormente no resulte óptimo. Como contrapartida, pue-

de optarse por implementar un clasificador kNN, que añade a los perfiles g_{it} una muestra de perfiles desplazados (esto es, hacia el interior y el exterior del objeto). Por tanto, la probabilidad p(b|g) de que un perfil recaiga en el contorno del objeto segmentado puede ser estimada escrutando los k vecinos más cercanos de g y calculando la tasa de perfiles correctamente asignados $b_k(g)$ a lo largo de ellos. Para evitar probabilidades nulas, puede ser utilizado un clasificador kNN moderado y definir:

$$p(b|g) = \frac{b_k(g) + 1}{k+2}$$
(3.15)

Modelo deformable

El modelo deformable se define como una malla triangular M = (V, E)de vértices $p, q \in V$ y bordes $[p,q] \in E$. M tiene la misma topología que el modelo estadístico de forma, esto es, para cada punto singular p en la malla, existe un correspondiente \tilde{p} en el modelo estadístico. La evolución del modelo deformable está controlada por la ecuación langragiana del movimiento: para cada punto singular p_i existe una fuerza reguladora interna $F_{int}(p_i)$ y una fuerza externa controladora de datos $F_{ext}(p_i)$. Estudiemos en detalle la naturaleza de estas fuerzas.

Fuerzas Internas Las fuerzas internas deben mantener la forma del modelo deformable similar a la del correspondiente modelo estadístico. Se define la similitud evaluando diferencias en la longitud de los bordes y en los ángulos formados entre caras vecinas. Esta aproximación está basada en los conceptos de tensión y rigidez, que también son los utilizados para definir la energía interna en la segmentación *snake*.

Para implementar la fuerza de tensión, cada borde [p, q] se modela como un resorte con elongación natural $|\tilde{p} - \tilde{q}|$. Por tanto, la fuerza de tensión en un punto singular p en la dirección que va hacia q es:

$$F_T(p,q) = \alpha \left(1 - \frac{|\tilde{p} - \tilde{q}|}{|p - q|}\right)(p - q)$$
(3.16)

donde α define la intensidad de la fuerza de tensión y es constante para todo borde $[p,q] \in E$. Por tanto, la fuerza de tensión total para un punto singular es la suma de todas las fuerzas a lo largo de todos sus bordes:

$$F_T(p) = \sum_{[p,q]\in E} F_T(p,q)$$
 (3.17)

Para describir las fuerzas de rigidez, haremos uso de las siguientes definiciones: para cada borde $[p_1, p_2] \in E$, los triángulos adyacentes $[p_1, p_2, q_1]$ y $[p_2, p_1, q_2]$ forman un ángulo θ correspondiente al ángulo del modelo estadístico. Se denomina a q_1, q_2 vértices exteriores de $[p_1, p_2]$ y a su vez forman el conjunto $V_O[p_1, p_2]$. La fuerza de rigidez para un punto singular exterior $q \in V_O[p_1, p_2]$ se define como:

$$F_R(q, [p_1, p_2]) = T(q, [p_1, p_2], \beta \delta) - q$$
(3.18)

donde $T(q, [p_1, p_2], \phi)$ es una rotación de ϕ grados del punto q alrededor del borde $[p_1, p_2]$, y β es la intensidad de la fuerza de rigidez. Para definir δ para ambos vértices externos, debemos considerar el caso de una constelación en la que la distancia d_1 entre $[p_1, p_2]$ y q_1 es distinta de d_2 . Para hacer el balance de fuerzas en ambos vértices exteriores, el triángulo con la menor distancia dtiene que rotar más que el otro, esto es:

$$\delta = \frac{d_{op}}{d_1 + d_2} (\theta - \overline{\theta}) \tag{3.19}$$

donde d_{op} es la distancia entre $[p_1, p_2]$ y el vértice contrario.



Figura 3.9: Dos triángulos adyacentes forman el ángulo θ . Las fuerzas de rigidez afectan a los vértices q_1, q_2 , pero también afectan a los puntos p_1, p_2 a mantener el equilibrio de la constelación.

Hay que tener en cuenta que las fuerzas internas no pueden alterar la posición global de la superficie deformable y, por tanto, la suma total ha de ser cero. Esto es, las fuerzas sobre q_1 y q_2 deben ser neutralizadas. Esta fuerza neutralizadora es la misma en los cuatro vértices de la constelación y se define como:

$$F_N([p_1, p_2]) = -\frac{1}{4} \Big(F_R(q_1, [p_1, p_2]) + F_R(q_2, [p_1, p_2]) \Big)$$
(3.20)

Finalmente, la fuerza de rigidez sigue por tanto la siguiente expresión:

$$F_{R}(p) = \sum_{[p_{1},p_{2}]\in E} \begin{cases} F_{R}(p,[p_{1},p_{2}]) + F_{N}([p_{1},p_{2}]) & \text{si } p \in V_{O}([p_{1},p_{2}]) \\ F_{N}([p_{1},p_{2}]) & \text{si } p = p_{1} \lor p = p_{2} \\ 0 & \text{en otro caso} \end{cases}$$
(3.21)

Por último, definimos la fuerza interna de un punto singular dado como la suma de las fuerzas de tensión y rigidez, esto es:

$$F_{int}(p) = F_T(p) + F_R(p)$$
 (3.22)

Fuerzas Externas Las fuerzas externas llevan la superficie deformable hasta la posición que más se ajusta a los datos de que se dispone. La bondad de este ajuste se evalúa usando modelos de apariencia local para todos los puntos singulares $p \in V$. Y no sólo en p, sino también en K posiciones adicionales a ambos lados de la superficie, lo que nos lleva a 2K + 1 pruebas para cada punto singular. Si definimos la posición de prueba óptima como s(p), una fuerza de elongación lineal dirige el punto singular a la dirección correspondiente de la siguiente forma:

$$F_{ext}(p) = \gamma(s(p) - p) \tag{3.23}$$

donde γ es la intensidad de la fuerza externa.

Usualmente s(p) puede ser determinado independientemente para cada punto singular p, pero el coste computacional asociado es intolerable. En la literatura existen diversos estudios para reducir el coste computacional asociado a esta operación, en la que nos detendremos si es éste el algoritmo elegido finalmente.

3.3.4.1. Crecimiento de regiones por mapeado de conectividad borrosa

Dehmeshki et al. [11] utilizan una técnica basada en crecimiento de regiones y lógica borrosa para segmentar tejidos blandos. Hemos escogido este trabajo por ajustarse en una primera fase al algoritmo de segmentación genérico propuesto por el Grupo de Tratamiento de la Señal y Comunicaciones para la herramienta VirSSPA, que podría respetarse, y al que se podría incluir una segunda fase basada en conectividad borrosa para la segmentación de tejidos blandos.

El diagrama de flujo del algoritmo se muestra en la figura 3.10.



Figura 3.10: Diagrama de flujo del algoritmo de segmentación basado en crecimiento de regiones por mapa de conectividad borrosa.

Obtención de semilla

La semilla, en el interior del objeto a segmentar, puede ser dispuesta por el facultativo.

Búsqueda de la Región de Interés

Para localizar una primera aproximación de la región de interés podemos utilizar, en primer lugar, una segmentación adaptativa local, y, posteriormente, un procedimiento de llenado de huecos, tal y como describimos a continuación.

 Segmentación Adaptativa Local. Un primer paso de segmentación basado en niveles de gris y adaptativo, por ejemplo, como el propuesto por el grupo de Tratamiento de la Señal, podría diferenciar entre tejido graso y otros tejidos de mayor nivel de gris aun en imágenes de bajo contraste (aunque nosotros nos aseguramos, con el paso de preprocesado, un buen contraste para nuestras imágenes).

2. Llenado de huecos. En imágenes de bajo nivel de contraste y sin presencia de tejido óseo, esto es, imágenes en las que varían mucho los niveles de gris dentro de un intervalo muy pequeño, la segmentación adaptativa local puede dejar huecos asignados al fondo aunque verdaderamente pertenecen a la región de interés inicial. Es necesario detectar estos fallos e incluir esas áreas en la región de interés.

Creación de la Máscara

En primer lugar, se excluyen los objetos no conexos incluidos inicialmente en la región de interés y posteriormente se aumenta la región de interés hasta que se deja fuera, efectivamente, toda la zona correspondiente al fondo. La máscara será utilizada para obtener los parámetros iniciales del crecimiento de regiones, además de como condición de confinamiento del mismo. También será utilizada para obtener la semilla óptima que hace al algoritmo reproducible.

Construcción del mapa de conectividad borrosa

El mapa de conectividad borrosa responde a la técnica de crecimiento de regiones del mismo nombre. La fuerza de conectividad borrosa se determina adaptativamente gracias al crecimiento de regiones a partir de una semilla hasta cubrir todos los vóxeles indicados por la máscara. Los parámetros requeridos para este paso se calculan haciendo uso de la *expectación máxima modificada* a través de la máscara, que responde a un modelo estadístico descrito ampliamente en el artículo, y al que nos remitiremos en caso de que ésta sea la técnica seleccionada para el algoritmo definitivo.

Crecimiento de regiones de orientación esférica

Se aplica un algoritmo de crecimiento de regiones de orientación esférica al mapa de conectividad borrosa a través de la máscara. La extensión del volumen segmentado se establece como el contorno de máximo contraste dentro de la máscara, esto es, la región que está menos conexa al fondo y más conexa a la región de interés. El crecimiento de regiones es directamente proporcional al índice de conectividad indicado por el mapa de conectividad borrosa e inversamente proporcional a la distancia de cada punto del vecindario de la región crecida al centro de la misma, de manera que vóxeles cercanos al centro de la región son incluidos antes que vóxeles alejados del mismo.