

Contornos activos para segmentación en imágenes digitales

Contornos activos sobre imagen de capas de lípidos.



Jorge Jara

Estudiante, Programa de Doctorado en Ciencias mención Computación DCC, Universidad de Chile. Ingeniero Civil en Informática, Licenciado en Cs. de la Ingeniería mención Informática, Universidad Austral de Chile. Líneas de investigación: Procesamiento de Imágenes y Segmentación, Geometría Computacional.
jjara@dcc.uchile.cl

Los continuos avances y la masificación de la computación y la electrónica en décadas recientes han repercutido con fuerza en el desarrollo del procesamiento de imágenes y su impacto en diversos ámbitos del quehacer humano, aprovechando la información que es posible obtener a partir de una imagen. Sistemas satelitales, microscopía, cámaras de todo tipo, ultrasonido, resonancias magnéticas, entre muchas otras fuentes, generan un volumen ingente y en constante aumento de datos que sirven a procesos de registro, inspección y/o control visual, con propósitos que varían según el área de aplicación y aun con los objetivos o escenarios particulares de sus usuarios: robótica, entretenimiento, investigación científica, medicina, vigilancia y procesos de manufactura, por nombrar algunos.

Sistemas como procesadores digitales de señales con capacidades en constante mejora, métodos como agrupamiento (*clustering*) o reconocimiento de patrones, y

los sistemas de cómputo de uso general han permitido la aplicación de métodos de cada vez más alto nivel para abordar problemas de complejidad y tamaño creciente en el tiempo. Por ejemplo, los microscopios ópticos modernos pueden recolectar cientos de gigabytes en minutos para estudiar procesos biológicos in vivo, con decenas de imágenes por segundo; telescopios de última generación pueden producir imágenes ópticas o de radiofrecuencia del cielo, del orden de terabytes por minuto (como los proyectos en desarrollo para telescopios de radio y ópticos en el norte de Chile).

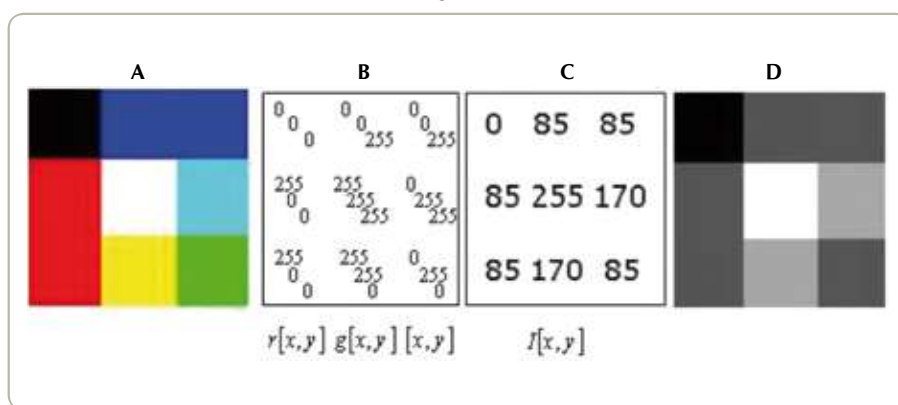
Tanto el volumen como la complejidad de los datos hacen necesario disponer de herramientas y modelos para su almacenamiento, tratamiento y análisis con algún nivel de automatización, constituyendo actualmente desafíos para matemáticas, computación, electrónica y disciplinas afines. A continuación se presentan los conceptos generales de los modelos de

contorno activo, que son métodos para modelar objetos de interés en imágenes, y que se encuentran dentro de las llamadas técnicas de segmentación.

ALGUNOS CONCEPTOS PREVIOS

Consideremos una *imagen* como una representación de uno o más objetos, generada por proyecciones basadas en las variaciones de la luz sobre ellos. En forma más general, es posible generar imágenes a partir de ultrasonido, resonancia magnética, temperatura u ondas de radio. Cualquier imagen en n dimensiones puede ser expresada como una función $I = I(x_1, x_2, \dots, x_n)$, cuyos valores son una característica observable o medible como intensidad de luz o color, profundidad, temperatura, densidad, etc. (un ejemplo muy simple se muestra en la fig. 1). Por ejemplo, una fotografía digital es una función de dos dimensiones (un plano) que podemos ver como una cuadrícula, en cuyos casilleros (píxeles) se registra el color de la escena fotografiada: la cámara que registra la imagen cuantifica el color de la escena y lo codifica con un valor numérico en cada píxel. Dispositivos más sofisticados son capaces de registrar imágenes tridimensionales de objetos en aplicaciones que van desde la microscopía atómica hasta vastas regiones del universo.

El *procesamiento de imágenes* se entiende como uno o más procesos de manipulación y/o tratamiento sobre un conjunto de imágenes de entrada para generar una



Formación de una imagen digital. Una imagen digital en escala de grises se define con una sola matriz $I[x, y]$. Cada posición de la matriz define un píxel y tiene valores en una escala que representan la intensidad I ; en escala de 8 bits, los valores entre 0 y $2^8 - 1 = 255$ codifican desde el negro hasta el blanco. Una forma para presentar colores en imágenes digitales imita al sistema visual humano: una imagen en color se puede definir combinando tres matrices o canales: $r[x, y]$ para el canal rojo (red), $g[x, y]$ para el canal verde (green), y $b[x, y]$ para el canal azul (blue). A: imagen digital en colores. B: representación r-g-b de los valores de intensidad para la imagen en A. C: imagen en escala de grises, con un solo valor de intensidad asociado (D). D: codificación en escala de grises. Fuente: www.scian.cl.

cierta salida. La diversidad y cantidad de procesos es tan amplia que se han distinguido tipos de procesos de acuerdo a su finalidad. Distingamos los siguientes tipos de procesamiento:

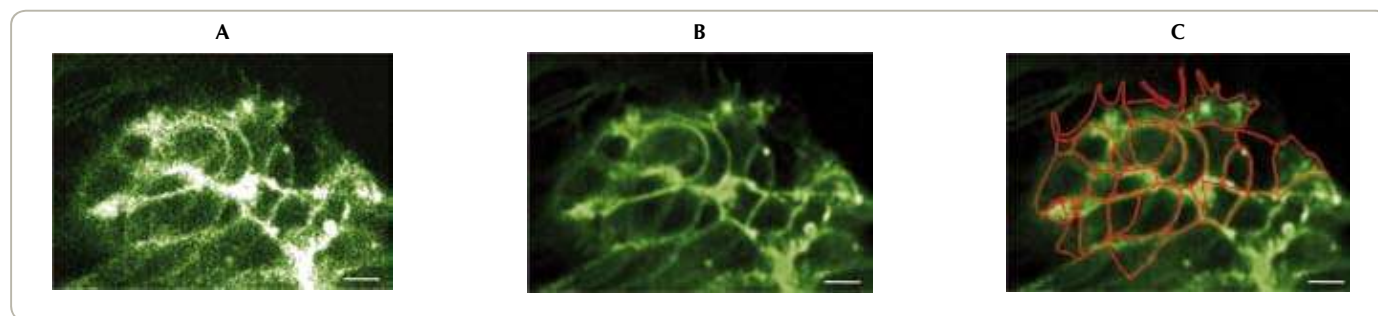
- Tratamiento de imágenes, que apunta a recrear una imagen lo más fiel posible a la realidad que representa, quitando o minimizando la influencia de elementos ajenos introducidos en la adquisición o almacenamiento de la imagen (como baja iluminación o ruido, por ejemplo).
- Análisis de imágenes, busca generar descripciones sobre una o más imágenes dadas, que den cuenta de características de interés.

- Comprensión de imágenes, a un nivel de abstracción mayor, busca interpretar y describir una o más imágenes en el contexto del problema que lo requiere.

Aunque esta clasificación es bastante general y algo simplificada, y existen otras que no necesariamente coinciden con ella, servirá como marco general para contextualizar el problema de segmentación como parte de una problemática más amplia y compleja.

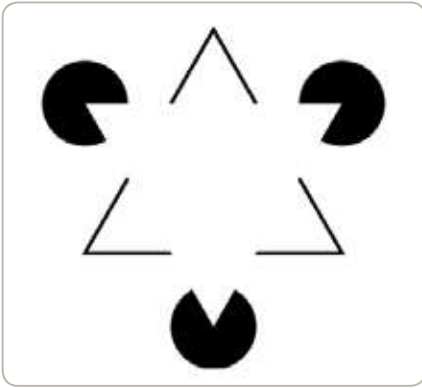
La fig. 2 muestra una imagen digital de células observadas mediante microscopía, que pasa por una etapa de tratamiento para minimizar el ruido y la distorsión del

Figura 2



Tratamiento y segmentación (análisis) en una imagen digital. A: imagen de microscopía de un conjunto de células cerebrales en un embrión de pez cebra, obtenidas mediante escáner de barrido láser. B: imagen tratada para minimizar la distorsión del proceso de adquisición empleando algoritmos de propósito específico. C: segmentación de bordes de estructuras celulares sobre la imagen; cada borde da cuenta de una región de interés, y su representación computacional permite acceso a descripciones geométricas de características como tamaño y forma. La barra de escala corresponde a $5 \mu\text{m}$. Fuente: SCIAN-Lab (datos no publicados).

Figura 3



Ejemplo de imagen con figuras subjetivas, los triángulos de Kanizsa.

microscopio, y que en una etapa de análisis es “descrita” mediante modelos de curva cerrada que demarcan las membranas en cada célula.

SEGMENTACIÓN

La segmentación es el proceso de subdividir una imagen en un cierto número de regiones, teniendo cada una de ellas alguna propiedad distintiva. Cuando se busca distinguir a un conjunto de objetos que presentan una propiedad específica, dichos objetos son denominados regiones de interés (*regions of interest, ROIs*): las personas en una fotografía o estrellas en una imagen de telescopio son ejemplos de ROIs. En general, las regiones son conjuntos disjuntos en la imagen, es decir que no se sobrelapan, de modo que cada elemento de la imagen (por ejemplo, un píxel en una imagen digital) pertenece a una sola región (se pueden hacer excepciones según el problema, cuando objetos se superponen o “comparten” una región del espacio, o bien no es posible resolver la imagen con más detalle para separarlos). El propósito de la segmentación es facilitar el acceso a información relativa a la imagen o a sus elementos constituyentes para etapas posteriores de análisis y extracción de información. El primer problema es que la definición misma de cada región es un problema que involucra aspectos psicológicos y físicos, (ver por ejemplo la fig. 3), de modo que no es posible establecer un criterio único u objetivo como *gold standard* para evaluar la pertinencia o calidad de una u otra técnica

de segmentación. La variedad de técnicas y formas de implementación existentes a la fecha, en combinación con lo anterior, obstaculizan la existencia de un enfoque unificado de segmentación, a pesar de lo cual existen trabajos en el área que apuntan a definir algunos criterios para cuantificar la calidad de distintos métodos.

La segmentación aparece como uno de los desafíos actuales en el procesamiento de imágenes a gran escala, como ocurre en biología y astronomía, en atención a requerimientos como la precisión y gran volumen de datos, objetos de morfología compleja y características particulares (por ejemplo, la fig. 2 que muestra entramados celulares y membranas con pequeñas prolongaciones a escala micrométrica). Métodos de segmentación han sido motivados específicamente por el análisis de imágenes en un campo particular, o suelen ser objeto de análisis y aplicación de técnicas nuevas, tanto por la masificación de tecnologías de computación como por sistemas ópticos que permiten observaciones a escala micro y nanométrica, muy cercana al nivel molecular capturando decenas o cientos de gigabytes de imágenes por experimento. Además, el análisis de objetos y complejos que cambian en el tiempo requiere tanto de modelos como implementaciones apropiadas y eficientes para su identificación y posterior descripción.

¿Cómo segmentar?

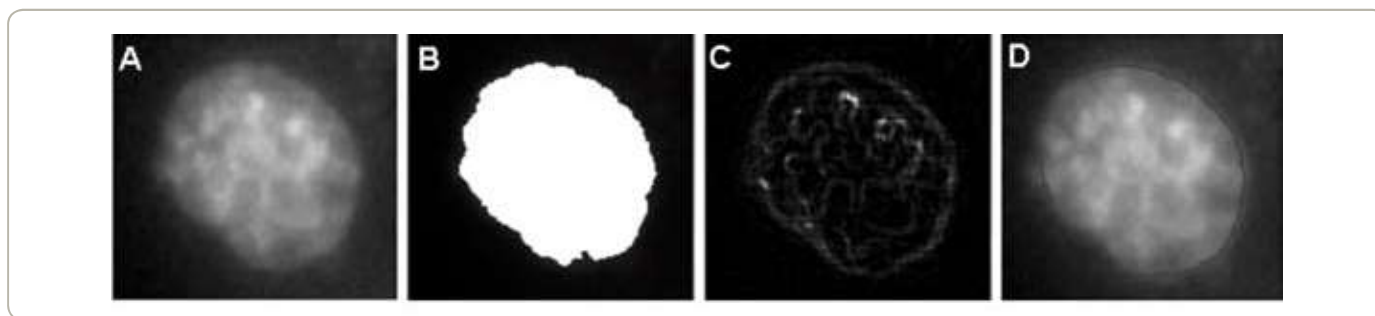
Con la premisa de que es de interés encontrar y describir ciertos objetos o regiones en una imagen, se puede aprovechar la forma en que la representamos: al considerarla como una función matemática, pueden evaluarse propiedades o características que permiten identificar o entregar pistas sobre las regiones de interés. Una forma directa es detectar **los cambios de color o intensidad** que existen en una imagen: la idea es que si cada objeto presenta un patrón de color o brillo relativamente homogéneo, se puede cuantificar la diferencia que se produce cuando se cambia de una región a otra (ver ejemplo en la fig. 4). En escenarios más complejos se pueden ajustar plantillas o patrones de formas, como por ejemplo probar

en qué sitio y de qué forma encaja mejor un rectángulo, una silueta de persona, etc. También se puede aprovechar información extra proporcionada junto con la imagen, como número de objetos, tamaños u otras características que “guíen” a la segmentación. Aprovechando las particularidades de distintas clases de imágenes, existe una gran variedad de modelos matemáticos y computacionales para segmentación, con diferentes grados de automatización y complejidad; ejemplos de esto lo constituyen modelos de *pattern matching*, *clustering*, redes neuronales, probabilísticos y funciones de optimización.

MODELOS DE CONTORNO ACTIVO

Los contornos activos son modelos de optimización que buscan balancear propiedades de imagen que definen a cada ROI junto con características del contorno de cada región (como por ejemplo su regularidad o curvatura), que son definidas a priori al formular el modelo de optimización: el contorno de cada región es modelado como una estructura elástica sometida a fuerzas que la deforman hasta que alcanza un estado de equilibrio – el óptimo – con mínima energía; por esta razón también se habla de “modelos deformables”. El estado óptimo se determina comúnmente mediante ecuaciones diferenciales que definen una condición de equilibrio entre distintas propiedades o *fuerzas*: por un lado se definen las fuerzas internas o propiedades intrínsecas a la forma del contorno, mientras que las fuerzas externas son las características de la imagen (como las transiciones de color o intensidad); también es posible definir fuerzas que mejoran la convergencia del ajuste o permiten completar información insuficiente en la imagen. La gran ventaja de los modelos de contorno activo es que admiten una amplia gama de formas, puesto que no condicionan una representación geométrica específica, sino que involucran varios grados de libertad que se ven restringidos por principios físicos que determinan su comportamiento según las propiedades definidas para las regiones a segmentar.

Figura 4



Segmentaciones por umbral y contornos activos en una imagen de escala de grises. A: imagen en escala de grises de un núcleo celular, obtenida por microscopía. B: Segmentación del núcleo (región de interés, en blanco) utilizando un umbral de intensidad, todos los píxeles con una intensidad mayor o igual al valor umbral son considerados parte del núcleo. C: imagen de gradientes de intensidad; se calcula la magnitud de los cambios de intensidad en la imagen original (A), en que los píxeles más brillantes corresponden a mayores gradientes o “saltos” de intensidad. D: segmentación del núcleo mediante contorno activo, utilizando como base el contorno de la región segmentada en B, y como fuerza de atracción la imagen de gradientes de C. Fuente: SCIAN-Lab (datos no publicados).

Los modelos de contorno han sido incluidos en formulaciones más generales, que permiten considerar otros elementos tales como las propiedades de la imagen que queda dentro y fuera de cada ROI, o el traslape de objetos. La extensión y complejidad del tema escapa al alcance de este artículo, por lo que se remite al lector interesado a la bibliografía.

Se distingue entre las formulaciones explícita e implícita, según la función de contorno, siendo representativos de cada una los *snakes* y los contornos activos basados en *level sets*, respectivamente, que se presentarán a continuación.

Si bien hasta ahora el modelamiento aparece más como un problema de corte más matemático que de “computines”, la conjugación matemático-computacional se da actualmente en equipos de trabajo y laboratorios de investigación de muchos países, con aplicaciones en robótica, astronomía, biología y medicina, por nombrar algunas. ¿La razón? Necesitamos adentrarnos un poco más en el tema...

MODELOS EXPLÍCITOS O PARAMÉTRICOS

Kass, Witkin y Terzopoulos [2] presentaron en 1988 un modelo para 2D que llamaron *snake* (serpiente): una curva en un plano es deformada por ecuaciones que pesan fuerzas internas y externas. Como fuerzas

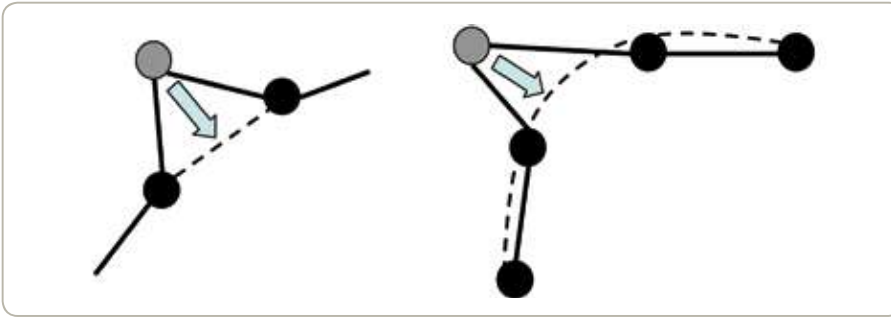
internas definieron elasticidad y rigidez, y como fuerza externa gradientes de intensidad de la imagen (fig. 5). Modelaron la curva como un conjunto finito de puntos $[x, y]$, representable en un programa de computador, con varios pasos de deformación hasta alcanzar un estado de reposo (óptimo), que definieron como el mínimo valor de la suma de las fuerzas evaluadas sobre todo el contorno, en forma de integral. La solución a esta minimización de la integral se puede obtener mediante ecuaciones diferenciales, definiendo ciertas restricciones y calculando las fuerzas en cada punto del contorno, deformándolo en iteraciones sucesivas desde un estado inicial. El estado inicial podía ser definido por el usuario, dibujando algunos puntos de control que permitían generar una interpolación para completar la curva. Además, Kass y cols. incluyeron fuerzas de atracción y repulsión definibles por el usuario, en forma de puntos específicos que llamaron de *resorte* y *volcán*, respectivamente; de este modo podían evitar que el *snake* se deslice hacia zonas no deseadas en la imagen.

MODELOS IMPLÍCITOS O GEOMÉTRICOS

El mismo año 1988, Osher y Sethian [3] presentaban un trabajo de física, en que modelaban la propagación de frentes; un anillo de fuego que se expande consumiendo un pastizal es un ejemplo. La expansión de

frente se modeló utilizando la curvatura y un tipo de ecuación que se popularizó por sus numerosas aplicaciones, una de las cuales se presentó más tarde como un modelo geométrico para contornos activos [4]. La idea general es que una curva cerrada en el plano se puede deformar con velocidad proporcional a su curvatura (mientras más plana se mueve más rápido, por ejemplo): si a esto se agregan fuerzas de imagen se obtiene un modelo parecido al de los *snakes*. Lo novedoso para este caso es el método que se emplea para resolver la ecuación: imaginemos primero que la curva de contorno 2D es parte de una superficie 3D que se deforma siguiendo una ecuación que incluye las reglas de evolución del contorno 2D; ahora pensemos que esta superficie es una especie de “mapa de elevación” en que la altura cero coincide con el contorno de la región, que se llama “curva de nivel cero” (de ahí el nombre de *level sets*, ver fig. 6); esta superficie se deforma siguiendo la ecuación de ajuste hasta que alcanza un estado de equilibrio, y se rescata el resultado buscando las curvas de nivel cero sobre la imagen. Este método permite que el número de contornos obtenidos sea independiente de con cuántas curvas empezó el ajuste. Mientras en el modelo de los *snakes* el número de ROIs debe coincidir con el número de curvas iniciales, para los *level sets* esto no es un problema, incluso es posible demostrar que su resultado (óptimo) no depende de cómo se inicialice el método.

Figura 5



Elasticidad y rigidez en un *snake*. En una curva de contorno, representada por un número finito de puntos, las propiedades o fuerzas de deformación (definidas en forma analítica) son calculadas aplicando una versión aproximada para cada punto, en función de sus puntos vecinos. Se muestra el efecto de calcular dichas fuerzas en un punto de prueba (gris). Izquierda: el efecto de fuerza elástica mueve al punto de prueba hacia la posición “promedio” entre sus vecinos. Derecha: el efecto de la fuerza de rigidez mueve al punto de prueba hacia la posición que corresponde a la curvatura estimada en sus vecinos.

EL ROL DE LA COMPUTACIÓN

El problema de contornos activos es un problema de optimización, y como tal se espera que su solución exista y se pueda aproximar por algún método, y es aquí donde entra la computación en juego. En la actualidad hay líneas de investigación activas en matemáticas al respecto, que buscan garantizar las condiciones de solución y convergencia de los métodos, así como formular propiedades que permitan segmentar imágenes con ruido, información faltante, oclusión de objetos, regiones con distintas propiedades, etc. Por su parte, el modelamiento y la implementación computacional en imágenes digitales que se valen de recursos algorítmicos, ha permitido aplicar diversos modelos de contorno activo con éxito en varias aplicaciones. A continuación veamos algunos ejemplos en que el uso de técnicas de algoritmos y geometría permite superar limitaciones de un modelo en el sentido matemático:

Snakes adaptativos. El método de los *snakes* es popular porque se formula y entiende con relativa facilidad, a pesar de que requiere de una buena inicialización, y la falta de una formulación que garantice una solución única. Para ambos casos han surgido algoritmos que permiten superar este tipo de obstáculos:

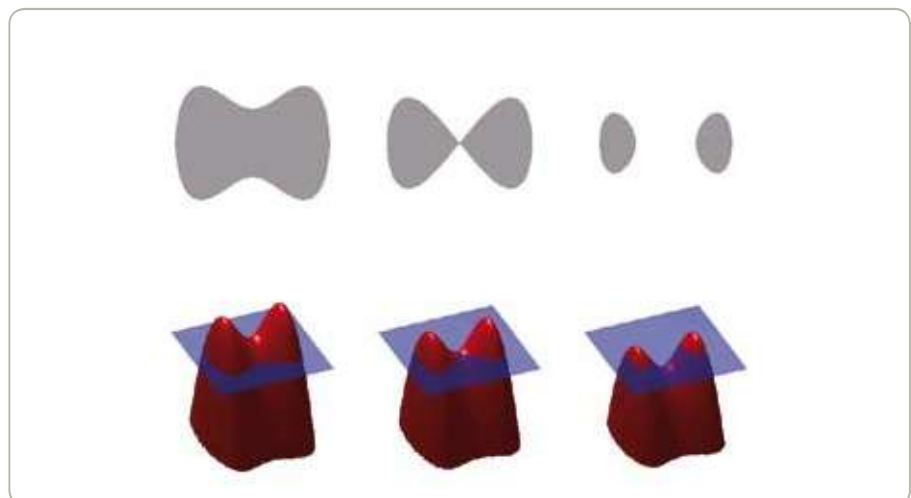
1. Si se inicializa una gran cantidad de contornos sobre la imagen y se sigue su evolución, se puede controlar cómo fusionar o dividir curvas: en el método de los *topology adaptive snakes* o *t-snakes* (publicado oficialmente años después de ser presentado en conferencia, [5]) se coloca una grilla sobre la imagen, que permite ir chequeando los cambios en el tiempo de los distintos *snakes*, y decidir si se eliminan, fusionan o separan curvas. De este modo es

posible utilizar una inicialización arbitraria, permitiendo incluso realizar segmentaciones automáticas cuando los parámetros del modelo son compatibles con las imágenes de entrada. A la fecha existen implementaciones tanto en 2D como en 3D.

2. La dependencia de la inicialización significa que un *snake* puede quedar “atrapado” en una zona de la imagen en que la suma de fuerzas sea cero pero que no corresponda a los bordes de una ROI, lo que vendría a corresponder con un óptimo local del problema de optimización. Cohen [6] presentó las fuerzas de inflación (que llamó *balloon forces*) con la siguiente idea: si un *snake* se encuentra en una zona que se considera interior de la ROI, se expande o “infla” (deformación en dirección normal hacia fuera), mientras que si esto ocurre en una zona considerada como exterior en la imagen el *snake* se contrae o “desinfla”.

Acelerando el cómputo de los level sets. Las ecuaciones de *level sets* se resuelven sobre la imagen completa como dominio, calculando una superficie iterativamente hasta que el resultado converge, para luego

Figura 6



Un ejemplo de curvas de nivel o *level sets* para contornos de formas en 2D. Arriba: tres regiones 2D en gris. Abajo: funciones de *level sets* de las regiones 2D; para cada región se tiene una función de superficie 3D, ϕ , cuyo valor cero (el contorno de cada región 2D) se muestra como la intersección con el plano de la imagen. Fuente: archivo de Wikimedia Commons, con licencia de dominio público.

realizar un posproceso de recuperación de las curvas de nivel. Pensemos en una imagen 2D compuesta por n píxeles en total. Con los métodos tradicionales de diferencias finitas o elementos finitos podemos requerir $O(n^3)$ operaciones para obtener la solución. Si los contornos (curvas de nivel cero) son muy pocos y/o pequeños hemos desperdiciado espacio y tiempo en el cálculo de porciones de la solución que no son de utilidad. Esta idea motiva el uso de algoritmos de cálculo de bandas (*narrowband level sets*), en que sólo se calcula la solución para una vecindad del nivel cero, la que se actualiza con cada iteración; también es posible hacer uso de técnicas de compresión para representar regiones homogéneas y usar estructuras de datos para acelerar las operaciones de cálculo y actualización de la superficie (usando el llamado *run length encoding*, [7]).

EN PERSPECTIVA

Hemos descrito de modo bastante general cómo el problema de segmentación de regiones de interés se aborda formulando modelos de contorno activo, que confieren a cada ROI ciertas propiedades que condicionan a la solución, y visto algunos ejemplos de cómo su implementación y aplicaciones conjugan aspectos de formulación-diseño en matemáticas y computación. La idea de fondo es que si es posible formular adecuadamente un modelo de optimización con las propiedades requeridas, se puede encontrar (o acercarse) una solución siguiendo un proceso de aproximaciones sucesivas hacia el óptimo. Este proceso puede “ayudarse” con técnicas complementarias: operaciones comunes con sistemas de ecuaciones algebraicas y diferenciales requiere de métodos numéricos que garanticen convergencia y a la vez sean eficientes y tan rápidos como se pueda; el uso de técnicas de tratamiento de imágenes para resaltar características o regiones y

mejorar el ajuste de los contornos hacia zonas de borde constituye un tema por sí solo; también el aumento de resolución de las imágenes y/o contornos permite mejorar significativamente la precisión al momento de describir características de morfología como volumen, superficie o complejidad de formas; en problemas de tracking o seguimiento de objetos se emplean técnicas de segmentación que toman en cuenta movimiento y elementos como coherencia entre imágenes para aumentar el nivel de información disponible y así mejorar los resultados. En años recientes, la disponibilidad de recursos de cómputo como procesadores de múltiples núcleos o GPUs permiten realizar segmentación y seguimiento en lo que un usuario percibe como “tiempo real”, o bien automatizar y manejar eficientemente el procesamiento en escalas de terabytes o superiores, constituyendo áreas de investigación y desarrollo en la actualidad.

SOBRE EL AUTOR

Jorge Jara es alumno del Programa de Doctorado del Departamento de Ciencias de la Computación (DCC) de la Universidad de Chile, y forma parte del laboratorio de procesamiento de imágenes científicas SCIAN-Lab (www.scian.cl) a cargo del Dr. Steffen Härtel en la Facultad de Medicina. SCIAN-Lab reúne a un grupo en la interface de computación, matemáticas e investigación biomédica, y colabora estrechamente con laboratorios del CMM y el DCC formando la Advanced Imaging and Bioinformatics Initiative AIBI (www.aibi.cl).

Jorge Jara y la profesora Nancy Hitschfeld del DCC de la U. de Chile, participan como colaborador y co-investigadora en el Proyecto Fondecyt 1090246: “Partial Differential Equations for 3D Photon Denoising, Optical Flow and Adjacent Active Surface Models

for High Throughput in Vivo Spinning Disk Microscopy”, que busca desarrollar técnicas de procesamiento de imágenes para microscopía con aplicaciones biológicas.

Además forman parte del Núcleo Milenio de Morfogénesis Neuronal (NEMO, www.nemolab.cl), una iniciativa de colaboración única en Chile y América del Sur que combina experticias en neurociencia molecular, biología del desarrollo, morfogénesis, neuropatología y cuantificación de imágenes in vivo para abordar las bases genéticas de la forma, estructura y organización funcional en el sistema nervioso central. BITS

REFERENCIAS

- [1] I. Young, J. Gerbrands, and L. van Vliet. Fundamentals of Image Processing. Delft: PH publications, 1995.
- [2] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *International Journal of Computer Vision* 1:321-331, 1988.
- [3] S. Osher and J.A. Sethian. Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulation. *Journal of Computational Physics* 79: 12-49, 1988.
- [4] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours. *Numerische Mathematik* 66: 1-31, 1993.
- [5] L. Cohen. On active contour models and balloons. *Computer Vision, Graphics and Image Processing: Image Understanding* 53(2): 211-218, 1991.
- [6] T. McInerney, D Terzopoulos. T-snakes: Topologically adaptable snakes. *Medical Image Analysis* 4(2): 73-91, 2000.
- [7] B. Houston, M.B. Nielsen, C. Batty, O. Nilsson, and K. Museth. Hierarchical RLE level set: A compact and versatile deformable surface representation. *ACM Transactions on Graphics*, 25(1):151-175, 2006.