

SEGMENTACION AUTONOMA Y EXTRACCION DE CARACTERISTICAS GEOMÉTRICAS EN DIATOMEAS

Enrique Alegre

Dpto. de Ingeniería Eléctrica y Electrónica. Universidad de León
dieeag@unileon.es

Saúl Blanco

Dpto. de Ecología. Universidad de León
dfqjbg@unileon.es

Manuel Castejón

Dpto. de Ingeniería Eléctrica y Electrónica. Universidad de León
dierfd@unileon.es

Raúl López

Escuela de Ingenierías Industrial e Informática. Universidad de León.

Lidia Sánchez

Dpto. de Ingeniería Eléctrica y Electrónica. Universidad de León
dielsg@unileon.es

Resumen

En este trabajo se ha realizado una aplicación que extrae automáticamente las principales características geométricas de las algas diatomeas. Estas algas resultan de gran utilidad en la monitorización de la calidad de las aguas, de ahí la importancia que la automatización de los procesos de análisis conlleva. La aplicación extrae 15 descriptores geométricos de cada diatomea almacenando su valor tanto en píxeles como en micrómetros. Se obtiene también la presencia o no del estigma al ser éste un valor determinante en la clasificación. Para llegar a este resultado es necesario realizar con anterioridad una segmentación autónoma de las imágenes digitales obtenidas mediante fotografía digital. Dicha segmentación autónoma requirió del desarrollo de un método específico basado en la combinación de dos técnicas: binarización por umbralización y detección de bordes utilizando el filtro de Canny. La separación de objetos solapados se resolvió mediante un método que combina binarizaciones, filtrados, adelgazamientos, operaciones morfológicas y el detector de bordes de Canny. Los valores geométricos calculados se compararon con los resultados proporcionados por un experto, tras realizar las mediciones de forma manual. Los resultados son muy similares por lo que se considera correcto el desarrollo realizado.

Palabras Clave: segmentación autónoma, solapamientos, descriptores geométricos, diatomeas

1. INTRODUCCIÓN

Las diatomeas o Bacilariofíceas son un grupo de algas unicelulares distribuidas en una gran variedad de ambientes acuáticos de todo el mundo. Hasta la actualidad se han descrito más de 10.000 especies diferentes, cada una de ellas adaptada a determinados rangos autoecológicos, de forma que se puede establecer una clara relación entre la composición de la comunidad de diatomeas y los parámetros fisicoquímicos del ambiente en que se desarrollan. Así, diversos trabajos publicados a partir de los años 70 del pasado siglo demuestran la eficacia de los métodos de control de la calidad del agua basados en el uso de la diatomeas como bioindicadores [1, 2, 3 4]. En España se están actualmente implementando en ríos y lagos redes de biomonitorización basadas en diatomeas, de acuerdo con los requerimientos de la Directiva Marco del Agua de la Unión Europea.

El principal factor limitante para la consolidación de las diatomeas como indicadores ecológicos radica en la dificultad de su identificación taxonómica. La aplicación de los índices diatomológicos más generalizados requiere frecuentemente diagnosis a nivel específico o subespecífico, lo cual implica importantes inversiones en equipos ópticos y en formación de expertos. Diversos ensayos de intercalibración [5, 6] han demostrado que los resultados de los índices biológicos basados en diatomeas son altamente sensibles al nivel de precisión en las determinaciones taxonómicas.

Los principales caracteres taxonómicos empleados en la identificación y clasificación de las diatomeas se

refieren a la morfología y ornamentación del frústulo. El frústulo es una cubierta silíceá formada por dos elementos (tecas) que encajan entre sí envolviendo a la célula. En la mayoría de las diatomeas pennadas, cada teca está recorrida longitudinalmente por un surco (rafe) dividido en dos ramas por un área central, donde ocasionalmente aparecen una o varias hendiduras (estigmas). Perpendicularmente al rafe se disponen numerosas estrías formadas por la alineación de varios poros o areolas.

En la actualidad, la identificación y conteo rutinario de diatomeas para el cálculo de índices biológicos se realiza mediante equipos de microscopía óptica de campo claro y de contraste de fase. La introducción de la microfotografía digital de alta resolución ha permitido el registro sistemático de gran cantidad de imágenes de diatomeas en visión frontal o valvar; lo que posibilita la creación de ficotecas electrónicas para la sistematización y estandarización taxonómica [7]. La diagnosis taxonómica clásica se realiza mediante claves de identificación o por comparación visual respecto a preparaciones tipo o a iconografías de referencia. En la actualidad, sin embargo, existen diversos estudios sobre la posibilidad de automatizar el proceso de identificación mediante análisis de imagen. Por ejemplo, el proyecto ADIAC (Identificación y Clasificación Automática de Diatomeas), de la Universidad de Algarve, ha obtenido resultados satisfactorios mediante la aplicación de herramientas de procesamiento de imágenes de diatomeas para la identificación del contorno y la ornamentación de cada frústulo y su posterior identificación [8]. Las técnicas de geometría morfométrica aplicadas sobre imágenes fotográficas de diatomeas también ofrecen resultados prometedores [9, 10]. El escaneado y enfoque automático de preparaciones microscópicas se encuentran en fases de desarrollo avanzadas, por el contrario la identificación de microestructuras morfológicas cercanas al límite de resolución de la imagen y la discriminación de frústulos de otros elementos en la muestra (células fraccionadas, partículas minerales) son las principales dificultades que encuentran los métodos de identificación automática de diatomeas.

En este trabajo presentamos una técnica de segmentación autónoma y extracción de características geométricas en diatomeas mediante análisis de imagen asistido por ordenador. Se describe una aplicación diseñada para el reconocimiento del contorno del frústulo y la obtención de sus principales descriptores morfométricos.

2. MATERIALES Y MÉTODOS

2.1 APLICACIÓN DESARROLLADA

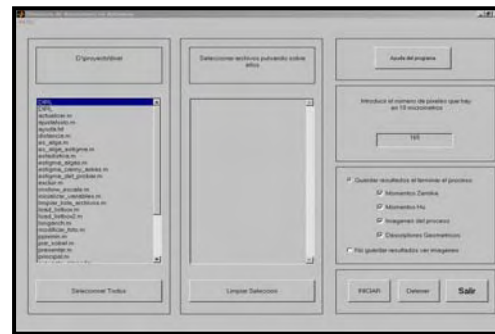


Figura 1. Selección de imágenes.

Se ha desarrollado una aplicación, utilizando Matlab y la Toolbox de Procesamiento de Imágenes que permite obtener automáticamente un conjunto de descriptores geométricos de diatomeas. La aplicación, Figura 1, está diseñada para procesar grandes cantidades de imágenes que se van seleccionando de la carpeta o carpetas en las que se encuentren.

Los resultados pueden almacenarse en archivos de texto delimitado, fácilmente importables por una hoja de cálculo o como una matriz de datos en Matlab. Es posible obtener descriptores Geométricos, o basados en los momentos de Hu y de Zernike.

Cuando se selecciona “Descriptores geométricos”, se

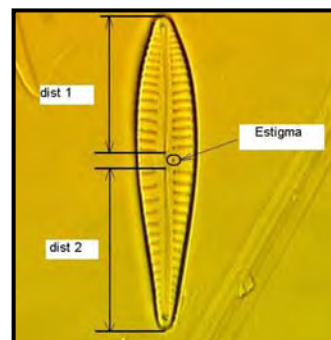


Figura 2. Estigma y distancias de rafe

crean dos archivos de texto en la misma ruta en la que se encontraba la última imagen seleccionada. El primer archivo contiene las medidas dimensionales obtenidas en píxeles y el segundo en micrómetros. Los valores en micrómetros dependerán de la ampliación con la que se haya obtenido la imagen. Se fija este valor estableciendo la relación del número de píxeles que hay cada 10 micrómetros.

Los descriptores geométricos obtenidos son:

- El área.
- Centro de masas coordenada X.
- Centro de masas coordenada Y.
- Eje mayor de la elipse equivalente con el mismo segundo momento.
- Eje menor de la elipse equivalente con el mismo segundo momento.
- Excentricidad.
- Área del más pequeño polígono que puede contener la imagen.
- Área del rectángulo más pequeño que puede contener la imagen.
- Diámetro equivalente.
- El resultado de la división del polígono más pequeño y el área total.
- El resultado de la división del rectángulo más pequeño y el área total.
- Longitud total.
- Anchura total.
- La distancia de rafe uno: dist_1 en la Figura 2.
- La distancia de rafe dos: dist 2 en la Figura 2.



Figura 3. Primera segmentación

Un valor determinante en la clasificación de las diatomeas es la presencia del estigma. Para cada alga se determina dicha presencia y se introduce como un nuevo valor en el descriptor codificando como 1 la presencia y 0 la ausencia.

En la Figura 2. puede observarse la posición del estigma y las distancias 1 y 2 de rafe.

Si se desea, puede indicarse que se guarden cuatro imágenes procedentes del procesamiento realizado: 1) una imagen en escala de grises del alga separada del fondo; 2) la imagen binaria resultado de la segmentación; 3) la imagen en escala de grises con el perímetro de la binaria extraído superpuesto; y 4) una

imagen en escala de grises con el perímetro del rafe y del estigma superpuesto, en caso de que exista.

2.2 SEGMENTACION Y EXTRACCION DE CARACTERÍSTICAS

La segmentación es un proceso no trivial debido a la aparición de artefactos de tamaño considerable, el solapamiento de diatomeas y la presencia de otras diatomeas menores o trozos de ellas en la imagen. Por ello ha sido necesario realizar esta operación en dos etapas.

2.2.1 Primera segmentación

Se extrae el alga o, si hay solapamiento, la región resultado de dicho solapamiento. Para ello se estudian las distintas regiones que aparecen en la imagen y se selecciona aquella que cumple un criterio de forma, obtenido de manera experimental, que indica una alta similitud con una diatomea. También se seleccionaría la región procedente de una diatomea solapada con otra.

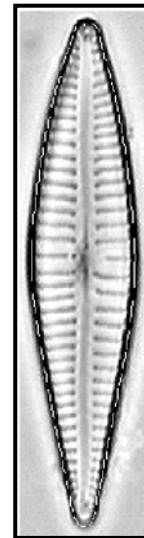


Figura 4. Segunda segmentación.

2.2.2 Segunda segmentación

Obtenido, mediante la primera segmentación, la región que contiene al alga, se procede a realizar una segunda segmentación más fina. El objetivo es determinar con precisión la ubicación de sus bordes y eliminar, cuando aparezcan, los objetos solapados.

En la Figura 4 se presenta el resultado de esta segunda segmentación. Se puede observar el resultado obtenido como una línea blanca discontinua superpuesta al borde de la diatomea. El método utilizado se explica en la sección 2.3.

2.3 DETECCIÓN Y ELIMINACIÓN DE OBJETOS SOLAPADOS

La separación de objetos superpuestos, en el caso de las diatomeas que no tienen ni forma ni tamaño constante, supuso desarrollar una técnica específica para el caso.

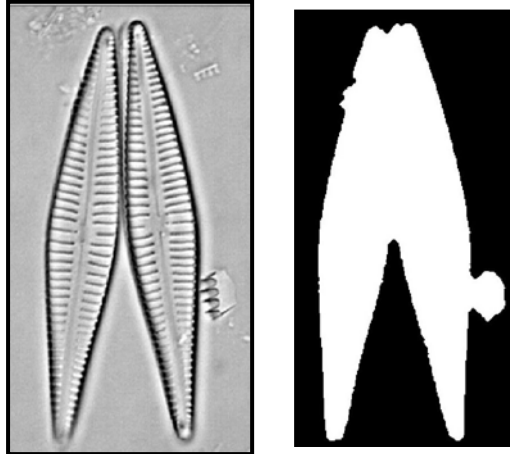


Figura 5. Diatomeas adyacentes y segmentadas.

Se utilizó una combinación de dos métodos de segmentación: binarización por umbralización y detección de bordes utilizando el filtro de Canny.

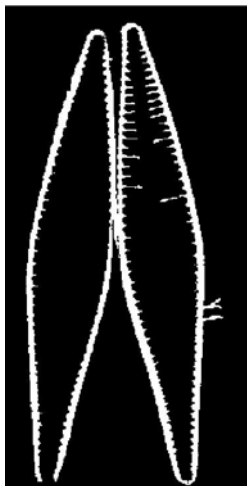


Figura 6. Negativo de la imagen binarizada.

La situación de partida es una imagen, similar a la que aparece en la Figura 5, que contiene dos diatomeas adyacentes. Al aplicar Canny para detectar

sus contornos y luego realizar una binarización a la imagen filtrada y una apertura, el resultado es el mostrado en la Figura 5 derecha. Puede verse que la adyacencia de las dos diatomeas y un tercer objeto producen una región mayor.

La obtención del contorno de los objetos solapados se realizó de la siguiente manera:

2.3.1 Binarización

Basándonos en que los bordes de las diatomeas son bastante oscuros en relación con el resto de la imagen, se realiza una binarización. El umbral se obtiene aplicando el método de Otsu.

2.3.2 Negativo

Se realizó un negativo de la imagen de manera que los bordes obtenidos anteriormente pasan a blanco y el resto a negro. En la Figura 6 se observa el resultado de esta operación. En este ejemplo el contorno ha quedado completamente definido, no siendo habitual obtener un resultado tan limpio y con el borde totalmente delimitado, aunque, para la detección del solapamiento, tampoco es necesario obtenerlo.

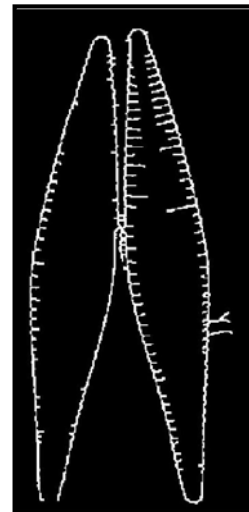


Figura 7. Adelgazamiento.

2.3.3 Filtrado

Para facilitar el posterior procesado se utiliza un filtro adaptativo de eliminación de ruido. Se empleó un filtro de Wiener para dos dimensiones. Aunque también se ha experimentado con otros filtros paso bajo, como es el filtro de mediana, los mejores resultados se obtuvieron con el indicado en primer lugar.

2.3.4 Adelgazamiento

A continuación se procede a reducir el grosor de los bordes anteriormente obtenidos. Para ello se realiza un adelgazamiento de la imagen anteriormente filtrada. Se pretende obtener los límites de cada objeto para así poder separarlo, ya estén los objetos solapados o sean adyacentes como en el ejemplo mostrado.

2.3.5 Dilatación

Como puede observarse en la Figura 7, el adelgazamiento puede provocar que se pierda la continuidad del contorno. Para obtener un borde conexo se procede a dilatar la imagen utilizando elementos estructurantes con forma de línea. En dicha figura se observa que el objeto pequeño que estaba tocando la diatomea de la derecha se ha convertido dos líneas por lo que ahora resulta sencillo de eliminar.

2.3.6 Canny y combinación

El paso final consiste en combinar la imagen binaria obtenida en el punto c) anterior, con el resultado, mostrado en la figura 5 derecha, de aplicar un filtro de Canny, con el postprocesado ya comentado, a la imagen original.

Tras realizar un complementario a la imagen procedente de la dilatación, mostrada en la figura 7, combinamos ésta con la de Canny mediante una operación AND. Puede verse el resultado en la figura 8 izquierda. En la figura 8 derecha se presenta el contorno del alga izquierda con los bordes en blanco obtenidos superpuestos. En las situaciones, como en esta, en las que nos quedarán dos objetos con forma de diatomea, se podrían haber calculado los descriptores de ambos, o bien, solamente de uno de ellos. La aplicación desarrollada descarta el objeto de menor tamaño y obtiene los descriptores del objeto mayor.

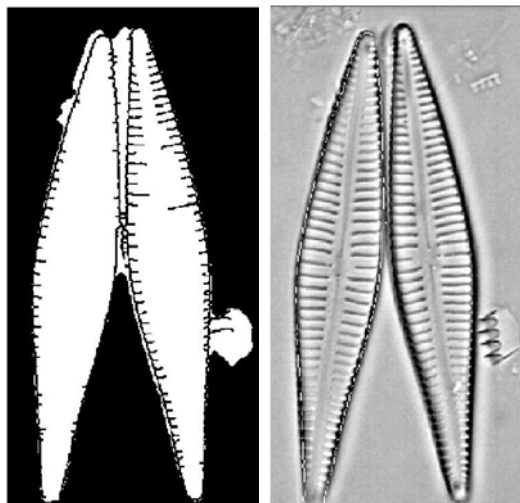


Figura 8. Imágenes con los objetos separados.

3. EXPERIMENTOS Y RESULTADOS

Algunos de los descriptores geométricos obtenidos mediante la aplicación desarrollada se compararon con los que se habían calculado midiendo sobre la imagen manualmente con una aplicación informática desarrollada para ese efecto. Los descriptores comprobados fueron la longitud y la anchura de las diatomeas, siendo estos considerados como dos de los valores fundamentales para caracterizar dichas algas. No se comprobó el resto al no disponer de datos supervisados para ello.

En la tabla uno pueden observarse estos resultados. La primera columna corresponde con el nombre de la imagen. En dicho nombre, los dos últimos valores, separados por una "x" corresponden con la longitud y la anchura de la diatomea.

Tabla 1. Muestra de experimentos

Nombre de la imagen	Longitud	Anchura
2 20x5	20,2482	6,0149
2 1 19 35x7	34,7887	7,4791
3 24x7	23,5201	7,2818
3 26x7	20,1132	8,0498
3 28x7	26,9097	7,3034
8 1 33x7	31,5151	6,7773
9 1 34x7	33,1517	7,5097
10 1 34x7	32,9099	7,4187
11 1 34x7	32,3641	7,9243
13 1 36x7,5	34,9093	7,5395
14 1 37x7	36,2424	7,7504
16 1 39x7	37,8199	8,0699
15 1 38x7	36,3636	7,2509
2 28x6	28,0000	6,5971
2 26x4	25,3339	5,4629
2 35x7	34,0034	7,1771
2 40x6	38,9102	6,9199
2 45x7	45,0988	7,4882
22 1 46x9	43,7683	8,3805
3 1 45x7	42,9152	8,7780

Puede observarse que los resultados obtenidos son bastante similares, si bien es cierto que no se han obtenido medidas cuantitativas de error.

Las pequeñas diferencias son debidas a errores en la determinación exacta de los bordes de la diatomea o,

en algunos casos, pudieran proceder de falta de precisión en la obtención manual de las medidas.

4. CONCLUSIONES

Se ha desarrollado con éxito una aplicación que realiza la segmentación autónoma de diatomeas y que mide las principales características geométricas que se utilizan en su clasificación. Los resultados obtenidos con la aplicación informática desarrollada fueron comparados con las mediciones realizadas por un experto. Se observó una gran similitud entre ambos resultados por lo que se concluye que tanto la segmentación autónoma como la extracción de características y la obtención de descriptores geométricos son correctas.

Con nuestra aplicación, el analista puede procesar grandes cantidades de microfotografías digitales, extrayendo de cada una de las algas el conjunto de descriptores geométricos que mejor describen su forma y más facilitan la posterior discriminación entre poblaciones.

El mayor logro de la aplicación es el ahorro de tiempo y esfuerzo necesario para realizar esta, de otro modo ardua, tarea. Asimismo, al contar con esta herramienta de extracción automática de características, podemos ya comenzar a pensar en el desarrollo de una metodología que permita un enfoque automático integral del proceso completo de monitorización.

En trabajos futuros se utilizarán otros tipos de descriptores para comprobar si es posible clasificar estas algas en función de ellos y sin utilizar las características geométricas ahora empleadas.

Referencias

- [1] Lange-Bertalot, H. (1979). "Pollution tolerance of diatoms as a criterion for water quality estimation". *Nova Hedw. Beih.* 64, pp. 285-304.
- [2] Whitton, BA., Rott, E., Friedrich, G. (eds) (1991). "Use of Algae for Monitoring Rivers". *Institut für Botanik, Universität Innsbruck, Innsbruck.*
- [3] Whitton, BA., Rott, E. (eds) (1996). "Use of Algae for Monitoring Rivers II". *Institut für Botanik, Universität Innsbruck, Innsbruck,*
- [4] Blanco, S., Ector, L., Bécares, E. (2004). "Epiphytic diatoms as water quality indicators in Spanish shallow lakes. *Vie Milieu* 53 (2/3)". *En prensa*
- [5] Prygiel, J., Coste, M., Ector, L. (1999). "Projets d'intercalibration européens et mise en place d'une charte de qualité diatomées pour l'IBD (et l'IPS)". *Cryptog. Algol.* 20, pp. 139-142.
- [6] Prygiel, J., Carpentier, P., Almeida, S., Coste, M., Druart, JC., Ector, L., Guillard, D., Honoré, MA., Iserentant, R., Ledeganck, P., Lalanne-Cassou, C., Lesniak, C., Mercier, I., Moncaut, P., Nazart, M., Nouchet, N., Peres, F., Peeters, V., Rimet, F., Rumeau, A., Sabater, S., Straub, F., Torrisi, M., Tudesque, L., Van de Vijver, B., Vidal, H., Vizinet, J., Zydek, N., (2002) . "Determination of the biological diatom index (IBD NF T 90-354): results of an intercomparison exercise." *J. Appl. Phycol.* 14, pp. 27-39.
- [7] Droop, SJM., Sims, PA., Mann, DG. and Pankhurst, RJ. (1993). "A taxonomic database and linked iconograph for diatoms". *Hydrobiologia* 269/270, pp. 503-508.
- [8] du Buf, H., Bayer. MM. (eds) (2002). "Automatic Diatom Identification". *Series in Machine Perception and Artificial Intelligence, Vol 51. World Scientific Publishing, London.*
- [9] Blanco, S., Fernández-Aláez, M., Bécares, E., Romo, S. and Ector, L. (2003), "Étude expérimentale sur l'interaction des communautés de diatomées épiphytes et du réseau trophique d'un lac peu profond (Lac de Sentiz, León, Espagne)". *Bull. Soc. Sci. Nat. Ouest. Fr. Suppl.* 2, pp. 288-304.
- [10] Blanco, S., Fernández-Aláez, M., Bécares, E., Romo, S. and Ector, L., (2003). "Variabilidad morfológica del género Gomphonema (Bacillariophyceae) en el epifiton de una laguna somera (León, NO España)". *XIV Simposio de Botánica Criptogámica.* Libro de resúmenes, pp. 23. Murcia. 17-19/12/2003.