

# TESELADO AUTOMÁTICO DE ORTOIMÁGENES

**G. Castilla Castellano**

IDR. Campus Universitario. 02071-ALBACETE (España). Correo electrónico: Guillermo.Castilla@uclm.es

## Resumen

Se presenta un método para obtener de forma automática, a partir de una ortoimagen, una malla de recintos irregulares de tamaño superior a un cierto mínimo. Esta malla es similar a la que delinearía un fotointérprete al que se le diera la tarea de dividir la imagen en regiones relativamente uniformes en su interior y diferentes de lo que las rodea, con la restricción de que todas superen un cierto tamaño mínimo. La malla puede usarse para elaboración o actualización de cartografía temática, para la estratificación de un inventario forestal, o para la división en rodales de un monte a ordenar. La malla constituye una plantilla inicial para el trabajo del fotointérprete, que básicamente sólo tiene que eliminar arcos en lugar de digitalizarlos, reduciéndose de esta forma tanto el tiempo como la subjetividad del trabajo. El formato de salida de la malla puede ser tanto ráster como vectorial, y los únicos parámetros que debe introducir el usuario son el tamaño medio deseado de los recintos, el tamaño mínimo exigido, y la precisión espacial de los bordes.

Palabras clave: *Segmentación de imágenes, Cartografía forestal, Raster, Vectorial*

## INTRODUCCIÓN

En la década pasada se propuso un enfoque alternativo de análisis del paisaje (WOODCOCK & HARWARD, 1992; WU & LOUCKS, 1995) que está adquiriendo cada vez mayor aceptación. Este nuevo paradigma concibe el territorio como un sistema jerárquico que posee una estructura vertical en niveles y una estructura horizontal en unidades anidadas u holones, que están compuestos de unidades de un nivel inferior y que a su vez forman parte de una unidad del nivel superior. La integración de cada holón es tal que las subunidades que lo componen (holones del nivel inmediatamente inferior) interactúan con más fuerza o con más frecuencia entre ellas que con subunidades de unidades vecinas, y ésa interacción es a su vez más débil que la que se produce en el seno de esas subunidades. Así, los

niveles superiores están constituidos por holones de mayor escala espacio-temporal y cuya integración es más débil, como ocurre en la jerarquía árbol-rodal-bosque.

Cada nivel se puede caracterizar por el tamaño medio de los holones que lo componen. Además, se puede asumir que por debajo de un cierto umbral de tamaño no existen holones de ese nivel. Ese umbral es análogo al tamaño de la unidad mínima cartografiable que se especifica en los mapas forestales (o de cobertura del suelo) para imponer un cierto nivel de generalización sobre el territorio. Entonces, si concebimos el mapa forestal como una representación de un cierto nivel de esa jerarquía, cada tesela de vegetación corresponderá a un holón, esto es, a un todo más o menos integrado que pueda verse como un objeto. Bajo esta perspectiva, los conceptos incluidos en la leyenda para describir las

teselas se refieren a objetos geográficos, que presentan una cierta integración entre sus partes y un cierto contraste con lo que les rodea. En cambio, en el enfoque tradicional empleado en teledetección, estos conceptos se suelen interpretar como referidos a materiales homogéneos que están constituidos por una serie de elementos (plantas) cuya naturaleza y patrón de distribución es recurrente, de forma que un trozo de terreno arbitrariamente delimitado dentro de la tesela sigue siendo un referente de ese concepto, y por tanto se pueden usar píxeles individuales como unidades básicas de análisis.

En el modelo jerárquico del paisaje, cada clase de cubierta se refiere a un conjunto de objetos u holones de un nivel particular que presentan una estructura y un funcionamiento similares. En consecuencia, las unidades básicas de análisis tienen a su vez que ser clasificables como objetos, y por tanto deben adaptarse a la variación espacial del fenómeno que se pretende clasificar. Dichas unidades las llamaremos rodales para diferenciarlas de las teselas del mapa (que normalmente serán agregaciones de las anteriores), y las definiremos como áreas simples (sin partes inconexas) que superan un cierto tamaño y cuyas especies dominantes, fisiognomía y/o fisiografía difieren perceptiblemente de las de sus alrededores. El método que se describe a continuación permite obtener dichas unidades básicas de forma automática partiendo de ortoimágenes (ya sean de una –pancromática– o varias bandas –multiespectrales o multitemporales–), y se basa en las siguientes hipótesis (CASTILLA, 2003):

1) *La variación espacial de la luminancia de la imagen coincide sensiblemente con la variación espacial de los atributos biofísicos relevantes de la cubierta.* Esto viene a decir que los bordes que identifiquemos en la imagen coincidirán con un cambio en la fisiografía, la fisiognomía y/o en la composición florística del estrato dominante de las regiones separadas por esos bordes. Aunque esto pueda parecerle una verdad de perogrullo a un fotointérprete, es necesario explicitar la relación entre los datos y el fenómeno estudiado cuando se usan métodos indirectos de análisis, máxime cuando la relación entre los números digitales de un píxel de una ortoimagen y los atributos biofísicos del trozo de

terreno correspondiente es variable y difícil de cuantificar. El que los bordes identificados en la imagen sean significativos dependerá de la conceptualización que se aplique al territorio, pero en cualquier caso servirán para la creación de unidades básicas a partir de las cuales construir objetos con significación ecológica. Los bordes que separan esos objetos son los lugares donde cambian las cualidades relevantes del territorio, donde las interacciones pasan de fuertes a débiles o de muy frecuentes a poco frecuentes.

- 2) *El nivel dentro del modelo jerárquico del paisaje en el cual se centra un estudio del territorio se puede expresar por el tamaño medio de los rodales que lo representan y por su tamaño mínimo.* Por tanto, para que un rodal pueda ser clasificado como una instancia –ejemplo particular– de un tipo de objeto de ese nivel, el rodal debe superar ese tamaño.
- 3) *Existe una correspondencia entre similitud semántica y parecido en la imagen de regiones adyacentes.* La agregación de regiones iniciales derivadas de la segmentación morfológica se basa en esta hipótesis.
- 4) *La complejidad de los bordes de los rodales aumenta indefinidamente con la escala de observación.* Esta hipótesis es un caso particular de la teoría de fractales (MANDELBROT, 1982), y en ella se funda la necesidad de trabajar con una resolución espacial adecuada al nivel de estudio del territorio. Un ejemplo de su validez se encuentra en la figura 5.

## METODOLOGÍA

Los parámetros de entrada del algoritmo de agregación de regiones limitada por tamaño (SCRM, de sus siglas inglesas Size-Constrained Region Merging) son: 1) tamaño medio en hectáreas de los rodales; 2) tamaño mínimo en hectáreas de los rodales; y 3) precisión de los bordes (en metros). La secuencia es la siguiente: SCRM ingiere la ortoimagen y eventualmente la remuestra a un tamaño de píxel adecuado a la precisión requerida; filtra la imagen mediante un suavizado con preservación de bordes; deriva la imagen de magnitud del gradiente de la imagen

filtrada; delimita el área de influencia de cada mínimo local del gradiente; agrega las regiones primarias hasta obtener una partición acorde a los parámetros de entrada; y convierte el ráster de salida en una capa vectorial.

### Remuestreo

El tercer parámetro de entrada, la precisión espacial de los bordes, se usa para definir el tamaño de píxel con el que se va a trabajar, y corresponde aproximadamente con la distancia media entre vértices de la capa vectorial de salida. Si la resolución de la imagen es mucho mayor que la precisión requerida, ésta se remuestrea a un tamaño de píxel igual a la mitad del valor de ese parámetro. De esta forma se obtienen bordes más acordes a la escala pretendida de representación, que de otra forma, en virtud de la hipótesis 4), resultarían excesivamente intrincados.

### Filtrado

Cualquier imagen de satélite, independientemente de su resolución, presentará zonas de textura gruesa que corresponden a objetos demasiado pequeños para que el sensor los pueda resolver. Como esos objetos en principio no interesan, se debe eliminar la variación que provocan en la imagen mediante un filtro que al mismo tiempo preserve los bordes de los objetos mayores. Esto se puede conseguir mediante un proceso de difusión iterativo, en el que la difusión es directamente proporcional al parecido entre píxeles vecinos. Durante este proceso, la imagen se va a autoorganizando en unidades formadas por grupos conectados de píxeles que cada vez son más parecidos entre sí, hasta que se llega a una estructura estacionaria en la que cada grupo constituye una unidad perceptiva en la imagen que denominaremos blob, esto es, una mancha homogénea, más oscura, más brillante o de un color diferente de lo que la rodea. De igual forma que los píxeles son los elementos básicos de la imagen desde el punto de vista físico, los *blobs* son las unidades mínimas perceptibles.

### Imagen de magnitud del gradiente

Cada blob tiene un píxel central que se puede asimilar desde el punto de vista perceptivo a un *atractor* local. El siguiente paso sería

localizar el área de influencia de ese atractor, o lo que es lo mismo, identificar el contorno de esa unidad perceptiva. Para ello es necesario crear una imagen de magnitud del gradiente que describa el valor absoluto de la máxima variación en similitud existente en el entorno de cada píxel de la ortoimagen filtrada. Si ésta es por ejemplo una ortofoto pancromática e imaginamos que corresponde a un modelo digital de elevaciones (DEM), la imagen de magnitud del gradiente sería simplemente el ráster de pendientes derivado de ese DEM. En el caso de imágenes a color o multiespectrales, la variación en similitud se mide a través de la distancia euclídea entre píxeles vecinos en el espacio multivariante definido por el conjunto de bandas de la imagen. Los mínimos locales (aquellos píxeles cuyo valor es inferior al de sus ocho vecinos inmediatos) de la imagen de magnitud del gradiente serían los atractores perceptivos de la imagen. Esta analogía asume que el área de influencia de cada mínimo local del gradiente corresponde a un blob.

### Partición en microcuencas

El siguiente paso consiste en delimitar el área de influencia de cada mínimo local de gradiente. Para ello se usa el algoritmo de las microcuencas o *watershed* (VINCENT & SOILLE, 1991), un método de segmentación morfológica que apenas se usa en teledetección y que sin embargo en visión artificial se considera una herramienta básica. Este algoritmo equipara la imagen de magnitud del gradiente con un DEM de una llanura lunar llena de cráteres contiguos de diferentes formas, tamaños y alturas. La tarea sería delimitar las crestas de esos cráteres, y para ello se simula una inmersión en la que cada mínimo local de gradiente es un manantial por donde fluye agua subterránea con un caudal tal que la cota de la lámina de agua que cubre las zonas sumergidas es la misma en todo el territorio. Entonces, en los lugares donde el agua procedente de manantiales distintos se juntaría, se construye un dique de un píxel de espesor y de altura ligeramente superior al punto más alto del territorio. Se procede así hasta que todo el territorio queda inmerso, y la red de diques resultante constituye la partición en microcuencas buscada, que es un ráster en el cual cada píxel tiene

como número digital (DN) la etiqueta numérica de la microcuenca a la que pertenece, y los píxeles de borde tienen su DN igual a cero.

### **Agregación de regiones**

Las regiones resultantes del paso anterior serán en general demasiado pequeñas como para tener significación como objetos del nivel estudiado. Por tanto los blobs tendrán que agregarse hasta formar regiones que superen el tamaño mínimo que se les supone a los objetos de interés. Para ello se extrae la firma digital de cada blob, que es la media de los DN de los píxeles que están dentro de él, de forma que la similitud entre regiones vecinas se evalúa por medio de la misma métrica que se empleó para producir la imagen de gradiente. Entonces en cada iteración, se identifica el par de regiones adyacentes con mayor parecido y se une, recalculando las relaciones de vecindad y el vecino más parecido a cada región de la zona afectada por la nueva unión. El proceso continúa de esta forma hasta que la suma de i) el número de regiones que superan actualmente el tamaño mínimo exigido; más ii) el resultado de dividir por el tamaño medio deseado (primer parámetro de entrada) el área ocupada por regiones de tamaño inferior al mínimo exigido; es menor que el resultado de dividir el área total de la imagen por el tamaño medio. De esta forma se garantiza que el tamaño medio final de las regiones estará próximo al especificado. A partir de este momento, se tiene en cuenta el tamaño de las regiones a la hora de agregarlas, de forma que el par candidato se une únicamente si al menos una de las dos regiones del par es más pequeña que el umbral. De esta manera, se forman primero regiones más homogéneas, a las que progresivamente se van incorporando regiones más pequeñas y de más alto contraste con su alrededor.

Cabe señalar que bien el tamaño medio o bien el mínimo se pueden dejar en blanco al introducir los parámetros, si se desea un resultado especial. Si no se especifica tamaño medio alguno, la partición resultante constará de regiones todas ellas cercanas al tamaño mínimo pero superándolo. Este tipo resultado es adecuado para la clasificación orientada a objetos, pero es excesivamente profuso si se pretende utilizar como plantilla para la fotointerpretación asistida.

Por el contrario, si no se especifica tamaño mínimo, se obtendrá una partición con regiones de tamaño dispar que exhiben tanto más contraste con sus vecinas cuanto mayor sea el tamaño medio especificado. En cualquier caso, debe tenerse en cuenta que el tamaño medio especificado debe ser bastante inferior al de las teselas finales del mapa, de forma que la agregación semántica, derivada bien de la clasificación orientada a objetos o bien del fotointérprete, pueda discurrir sin excesivos problemas debido a la heterogeneidad de las unidades a clasificar.

### **Vectorización**

Finalmente, si los rodales resultantes se van a usar para fotointerpretación asistida, la partición resultante tiene que pasarse de formato ráster a vector. Para ello se considera que los píxeles de borde (con DN=0) constituyen un área de transición entre rodales que se puede representar por su eje medio, de forma que los vértices iniciales de la capa vectorial estarán constituidos por los centros de estos píxeles, y cada arco estará limitado por dos nodos, es decir píxeles de borde que tienen cuatro o más vecinos de borde en lugar de dos. Entonces para cada arco, se toman cada tres vértices intermedios y se sustituyen por su centroide, y se suavizan las líneas que conectan cada centroide mediante una interpolación segmentaria cúbica o spline.

### **EJEMPLOS**

El algoritmo SCRM se ha implementado en IDL ([www.rsinc.com](http://www.rsinc.com)) y se ha aplicado en numerosas imágenes. Como ejemplo, se muestran unos resultados sobre imágenes multispectrales, pancromáticas y multitemporales. La figura 1 es un trozo de 5x5 km<sup>2</sup> de una imagen Landsat ETM+ (RGB 453) capturada el 22/08/2000 sobre una zona rural de Guadalajara. La figura 2 corresponde a un trozo de 2,5x2,5 km<sup>2</sup> de una ortofoto pancromática de 1 m de resolución tomada en abril del 97 sobre la misma zona. La figura 3 es un trozo de 5x5 km<sup>2</sup> de una imagen NDVI multitemporal (R=20/5/00; G=28/6/00; B=16/7/00) sobre una zona agrícola en Barrax. Finalmente, la figura 4 muestra un detalle de la delineación del borde de un quemado a dos reso-

luciones diferentes, 1m y 10m. El borde correspondiente a la resolución menor es mucho más intrincado el obtenido a 10m.

## DISCUSIÓN

La partición de salida de SCRM produce una buena impresión visual en todas las imágenes en que se ha probado. En la mayoría de los recintos, se aprecia una cierta coherencia interna y una cierta discontinuidad a lo largo del borde entre cada recinto y sus vecinos. Esta partición es un modelo del territorio que se genera mediante dos mecanismos separados de generalización: la regularización efectuada por el sensor (y más tarde por el filtro) y la agregación de unidades perceptivas según su similitud radiométrica. El método de segmentación que materializa el segundo mecanismo debería, si pretende ser universal, estar únicamente influenciado por los datos de la imagen, sin requerir más información del usuario que los parámetros de tamaño. Sin embargo, no está claro qué combinación de regularización y agregación es la mejor para obtener este modelo primario. Además, cambios en el tipo de muestreo, la intensidad del filtro y el criterio intermedio de parada (a partir del cual se tiene cuenta el tamaño de las regiones) dan lugar a resultados diferentes, particularmente en zonas de contraste bajo. No obstante, la falta de solución única también se manifiesta entre los fotointérpretes, cuyas interpretaciones de esas zonas son también susceptibles de diferir notablemente.

Esas inconsistencias aparecen a fin de cuentas porque la cobertura del suelo se puede representar de muy diversas maneras, y a menudo no se puede decir contundentemente que una de ellas sea la mejor, pues cada institución o individuo tienen diferentes intereses, conceptos y métodos. Además, la minimización de los errores de representación está en conflicto con el nivel de generalización del mapa, y en cualquier caso es dependiente de la propia definición de error. En definitiva, no existe una única estructura jerárquica anidada que se ajuste a un determinado paisaje (WIENS, 1995), puesto que las jerarquías y los mapas que las representan son artífices de los que no servimos para aprehen-

der la complejidad del territorio. Por tanto la bondad de un mapa de coberturas reside más en su utilidad que en su precisión, pues esta última es relativa a cómo se conceptualicen los hechos del territorio. El criterio decisivo es que el mapa permita al usuario hacerse una imagen mental de ese territorio y tomar decisiones sobre él que no difieran sustancialmente de las que respectivamente se hubiera hecho o habría tomado si lo conociera directamente.

## CONCLUSIONES

El método de teselado automático SCRM genera, a partir de una ortoimagen, una primera partición del territorio que bien puede usarse como plantilla inicial para un trabajo de fotointerpretación asistida, o bien para definir unidades básicas de una clasificación orientada a objetos. Comparado con otros métodos (BAATZ & SCHAPE, 2000) ya incluidos en software comercial, SCRM es menos costoso computacionalmente, necesita sólo tres parámetros del usuario que están relacionados explícitamente con el tamaño de los recintos de salida y con la precisión de sus bordes, y aborda la naturaleza fractal del paisaje y su estructura jerárquica de una forma conceptualmente coherente. En definitiva, SCRM proporciona un primer paso en la transformación de una representación numérica del territorio (la ortoimagen) en una representación simbólica orientada a objetos (el mapa de coberturas).

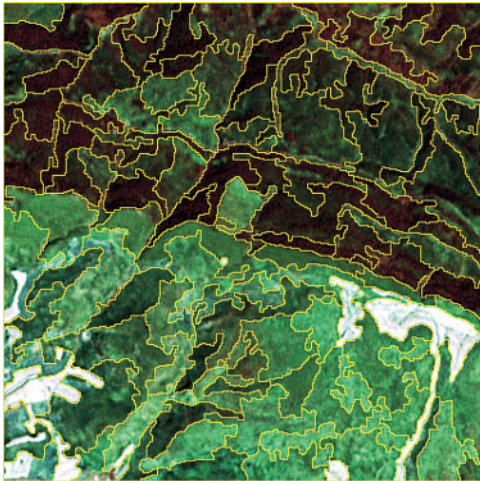
## BIBLIOGRAFÍA

- BAATZ, M. & SCHAPE, A.; 2000. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. *In: J. Strobl, & T. Blaschke (eds.), Ang. Geogr. Inf.* 12: 12-23. Wichmann-Verlag. Heidelberg.
- CASTILLA, G.; 2003. *Object-oriented analysis of Remote Sensing images for land cover mapping: conceptual foundations and a segmentation method to derive a baseline partition for classification*. Tesis doctoral. Universidad Politécnica de Madrid. Madrid.

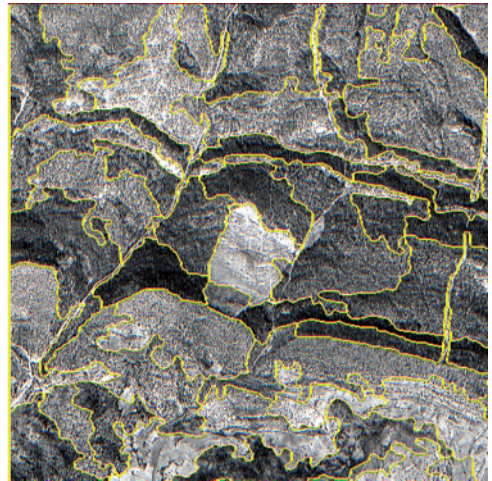
- MANDELBROT, B.B.; 1982. *The Fractal Geometry of Nature*. W.H. Freeman. New York.
- VINCENT, L. & SOILLE, P.; 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE PAMI* 13(6): 583-598.
- WIENS, J.A.; 1995. Landscape mosaics and ecological theory. In: L. Hansson, L. Fahrig & G. Merriam (eds.), *Mosaic Landscapes and Ecological Processes*: 1-26. Chapman and Hall. London.
- WOODCOCK, C.E. & HARWARD, V.J.; 1987. Nested-hierarchical scene models and image segmentation. *Int. J. Rem. Sens.* 13(16): 3167-3187.
- WU, J. & LOUCKS, O.L.; 1995. From balance-of-nature to hierarchical patch dynamics: a paradigm shift in ecology. *Quat. Rev. Biol.* 70: 439-466.

## TESELADO AUTOMÁTICO DE ORTOIMÁGENES

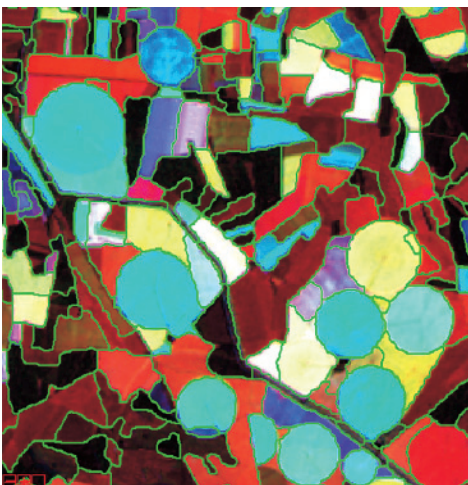
G. Castilla Castellano



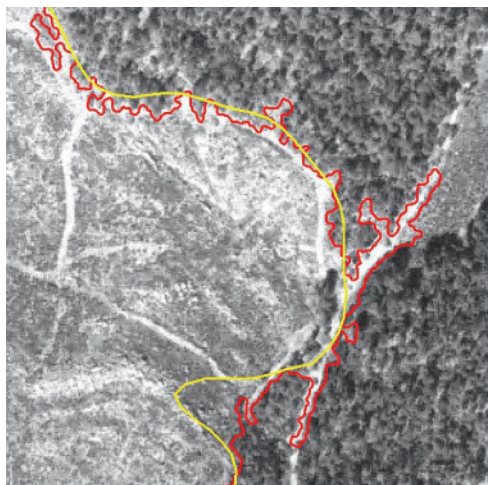
**Figura 1.** Teselado de un trozo de 5x5 km<sup>2</sup> de una imagen Landsat



**Figura 2.** Teselado de un trozo de 2.5x2.5 km<sup>2</sup> de una ortoimagen pancromática



**Figura 3.** Teselado de un trozo de 5x5 km<sup>2</sup> de una imagen NDVI multitemporal



**Figura 4.** Dos delineaciones del mismo rodal a diferentes resoluciones (1 y 10 m)