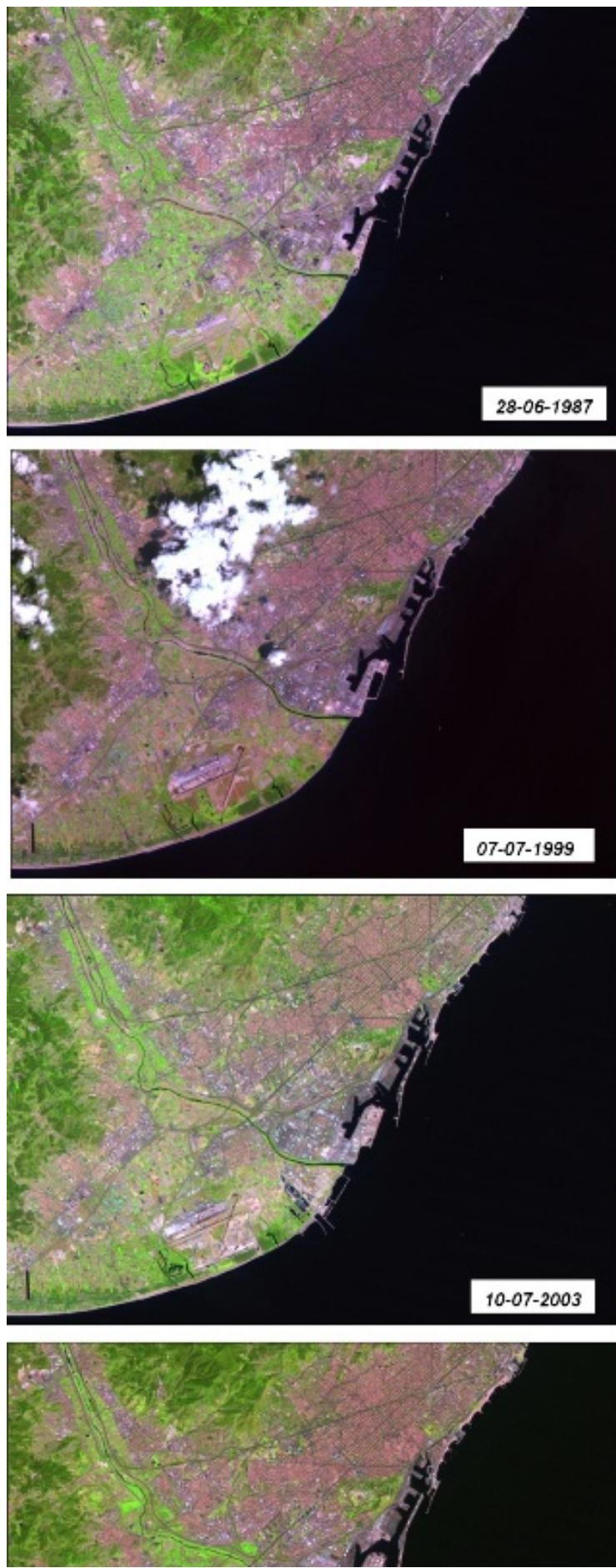




BARCELONA RECERCA I INNOVACIÓ

12/2014

Contribuciones al análisis automático de grandes volúmenes de datos





Esta tesis es fundamentalmente un estudio sobre la búsqueda de automatismos en diferentes algoritmos aplicados a la ciencia de la información geográfica. Los trabajos que la forman tienen un importante objetivo común: conseguir la máxima calidad posible de los resultados obtenidos al sustituir el método habitual (manual) en el análisis de grandes volúmenes de datos por el correspondiente método automático, que muestra varias ventajas respecto al primero, y en el que tienen un papel relevante los metadatos, es decir, la información sobre datos, objetos o recursos.

Los automatismos tienen varias ventajas respecto a un método manual: objetividad en la toma de decisiones, repetitividad, trazabilidad, constancia y homogeneidad (la calidad del proceso, la atención en términos humanos no decae por tareas repetitivas), y una muy importante en el hora de hacer investigación: podremos analizar las diferentes respuestas de nuestro modelo en base a repeticiones donde cambien unos pocos criterios o parámetros, podremos replicar a otros ámbitos o momentos en el tiempo, analizar diferentes escenarios, etc. Asimismo, desde el punto de vista operacional, añade una característica no viable por los métodos manuales: el tratamiento de un gran volumen de datos, el llamado *Big Data*. Este es un problema (bastante bienvenido, por cierto) de los nuevos tiempos y que va creciendo de forma acelerada. Actualmente se ha generalizado la libre distribución de datos y se generan con un detalle espacial y recurrencia temporal poco imaginables hace unos años.

Las principales aportaciones concretas siguiendo estos objetivos son:

- La automatización del análisis del variograma, el diseño distribuido y la implementación paralelizada de un método geoestadístico de interpolación espacial: el *kriging*. Permite usar este método en aplicaciones en tiempo real y con grandes volúmenes de datos, manteniendo la máxima calidad en el modelo interpolado.
- La integración en el SIG de todos los procesos para la publicación automática en Internet de mapas de variables continuas obtenidas desde muestras de diversas especies de flora y fauna a partir de modelizaciones geoestadísticas.
- El análisis geoestadístico de imágenes de Teledetección de largas series temporales ha permitido añadir elementos de estudio del patrón espacial que mejoran otras técnicas que no tienen en cuenta la estructura espacial de las propiedades radiométricas de las imágenes. Algunos ejemplos son la selección de imágenes de elevada calidad y el estudio

de las alteraciones provocadas por los métodos de compresión con pérdida.

- La generación automática de valores de referencia, en zonas con valores de reflectancia casi constante, y su aplicabilidad en un nuevo modelo de corrección radiométrica automática de imágenes Landsat, abre la puerta a estudios donde se necesite tratar un gran número de estas imágenes.

Finalmente, y como aportación global de la tesis, destacar el papel relevante que tienen los metadatos en los procesos automáticos. Este nuevo rol añade unos rasgos más geográficos que los usos habituales sólo aplicados a las búsquedas sobre catálogos, a menudo alejadas de su componente geográfica.

Imagen izquierda: 4 escenas, seleccionadas de la larga serie Landsat, que muestran los cambios de los últimos 30 años en la desembocadura del río Llobregat e infraestructuras vecinas, como el aeropuerto y el puerto de Barcelona.

Lluís Pesquer

l.pesquer@creaf.uab.cat

Referencias

"Contribucions a l'anàlisi automàtica de grans volums de dades en el context de la Ciència de la Informació Geogràfica", tesis doctoral de Lluís Pesquer (CREAF), dirigida por el Dr. Xavier Pons y leída en el Departamento de Geografía de la UAB.

[View low-bandwidth version](#)