

El modelo digital de elevaciones

por [Angel Manuel Felicísimo](http://www.etsimo.uniovi.es/~feli/), biólogo
<http://www.etsimo.uniovi.es/~feli/>



Introducción

En la cartografía convencional la descripción de las elevaciones a través del mapa topográfico constituye la infraestructura básica del resto de los mapas. El papel equivalente en los MDT lo desempeña el modelo digital de elevaciones (MDE), que describe la altimetría de una zona mediante un conjunto de cotas. Siguiendo la analogía cartográfica, es posible construir un conjunto de modelos derivados, elaborados a partir de la información contenida explícita o implícitamente en el MDE. Los modelos derivados más sencillos pueden construirse exclusivamente con la información del MDE y reflejan características morfológicas simples (pendiente, orientación, etc.). Incorporando información auxiliar es posible elaborar otros modelos más complejos, utilizando conjuntamente la descripción morfológica del terreno y simulaciones numéricas de procesos físicos.

Origen

El término *digital terrain model* tiene aparentemente su origen en el Laboratorio de Fotogrametría del Instituto de Tecnología de Massachussetts en la década de los años 50. En el trabajo pionero de Miller y Laflamme (1958) se establecen ya los primeros principios del uso de los modelos digitales para el tratamiento de problemas tecnológicos, científicos y militares. La definición del MDT que se menciona en sus trabajos es "*una representación estadística de la superficie continua del terreno, mediante un número elevado de puntos selectos con coordenadas (x, y, z) conocidas, en un sistema de coordenadas arbitrario*"

Puede observarse el uso del término MDT como sinónimo de MDE, así como el escaso énfasis en el aspecto de la referenciación geográfica —aunque en el trabajo, unas líneas más adelante, se indica la conveniencia de establecer una relación entre el sistema de coordenadas y un sistema de proyección geográfica—.

El objeto de su trabajo fue acelerar el diseño de carreteras mediante el tratamiento digital de datos del terreno adquiridos por fotogrametría, planteándose una serie de algoritmos para la obtención de pendientes, áreas, etc. El problema del número de datos se planteó de forma crítica, dada la escasa capacidad de almacenamiento de los ordenadores en aquella época, y una buena parte del esfuerzo del proyecto se dedicó a desarrollar métodos de representar los perfiles topográficos de la forma más sintética posible.

Las aplicaciones informáticas de uso menos aplicado tuvieron que esperar algunos años más y, habitualmente, surgieron en un "contexto SIG" que incorporaba la información topográfica para el manejo de cartografía digital en términos más generales. El *Harvard Laboratory for Computer*

Graphics and Spatial Analysis marcó un hito cuando presentó SYMAP en 1967. SYMAP era una aplicación formada por un conjunto de programas de manejo de cartografía digital que incluía algoritmos de interpolación para generar mapas de isopletas a partir de puntos acotados distribuidos irregularmente.

En los últimos años han surgido ya multitud de aplicaciones informáticas capaces de manejar eficazmente los MDT. A pesar de ello, aún queda pendiente, lo mismo que en los SIG, la resolución definitiva de problemas básicos como, por ejemplo, conseguir una estructura de datos idónea, conseguir algoritmos eficientes, o facilitar el uso de los sistemas por parte de los usuarios.

Definición y estructura del MDE

De acuerdo con la definición general presentada en el capítulo 1, **un modelo digital de elevaciones es una estructura numérica de datos que representa la distribución espacial de la altitud de la superficie del terreno.**

Un terreno real puede describirse de forma genérica como una función bivariable continua $z = \zeta(x, y)$ donde z representa la altitud del terreno en el punto de coordenadas (x, y) y ζ es una función que relaciona la variable con su localización geográfica. En un modelo digital de elevaciones se aplica la función anterior sobre un dominio espacial concreto, D . En consecuencia, un MDE puede describirse genéricamente como $MDE = (D, \zeta)$.

En la práctica, la función no es continua sino que se resuelve a intervalos discretos, por lo que el MDE está compuesto por un conjunto finito y explícito de elementos. Los valores de x e y suelen corresponder con las abscisas y ordenadas de un sistema de coordenadas plano, habitualmente un sistema de proyección cartográfica.

La **generalización** inherente a la discretización del modelo implica una pérdida de información que incrementa el error del MDE y, en consecuencia, se propaga a los modelos derivados. Por este motivo, se han ensayado numerosas opciones en la búsqueda de una forma de representar y almacenar la altitud que equilibre la pérdida de información y algunos efectos secundarios indeseables como el excesivo tamaño de los archivos o la dificultad de manejo.

Estructuras de datos en el MDE

De forma general, la unidad básica de información en un MDE es un punto acotado, definido como una terna compuesta por un valor de altitud, z , al que acompañan los valores correspondientes de x e y . Las variantes aparecen cuando estos datos elementales se organizan en estructuras que representan las relaciones espaciales y topológicas.

Mientras que los mapas impresos usan casi exclusivamente una única convención —las curvas de nivel— para la representación de la superficie del terreno, en los MDE se han utilizado alternativas algo más variadas. Históricamente, las estructuras de datos en los sistemas de información geográfica y, por extensión, en los modelos digitales del terreno, se han dividido en dos grupos en función de la concepción básica de la representación de los datos: **vectorial y raster**:

- **el modelo de datos vectorial está basado en entidades** u objetos geométricos definidos por las coordenadas de sus nodos y vértices
- **el modelo de datos raster está basado en localizaciones** espaciales, a cada una de las cuales se les asigna el valor de la variable para la unidad elemental de superficie

En el modelo vectorial los atributos del terreno se representan mediante puntos, líneas o polígonos con sus respectivos atributos. Los puntos se definen mediante un par de valores de coordenadas con un atributo de altitud, las líneas mediante un vector de puntos —de altitud única o no— y los polígonos mediante una agrupación de líneas.

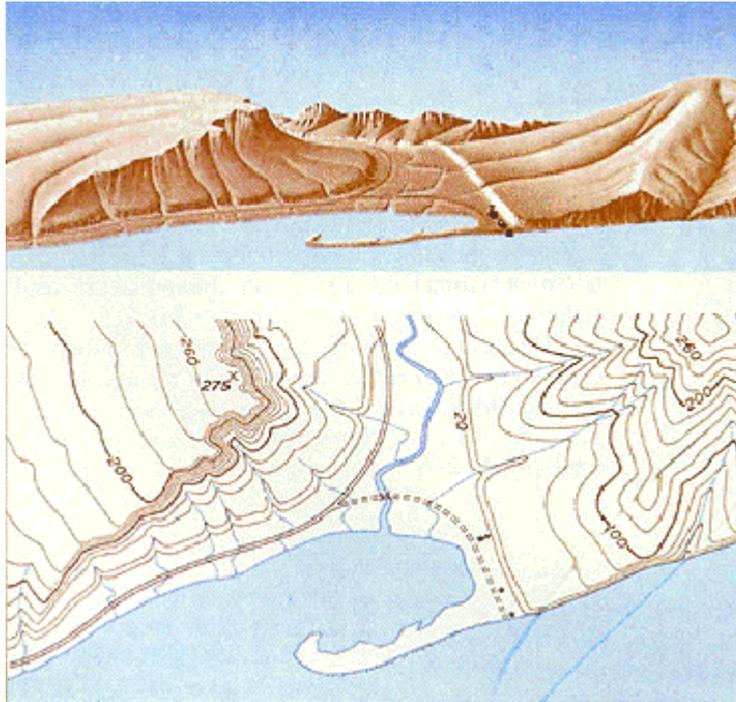
En el modelo raster, los datos se interpretan como el valor medio de unidades elementales de superficie no nula que teselan el terreno con una distribución regular, sin solapamiento y con recubrimiento total del área representada. estas unidades se llaman **celdas** o **teselas** y, si se admite la analogía con los términos usados en proceso de imágenes, *pixeles*.

Cada **modelo de datos** puede expresarse mediante diferentes **estructuras de datos**; dentro de los dos modelos básicos, la práctica y el tiempo han reducido las potenciales variantes de estructuración a unas pocas. Las más representativas son dos estructuras vectoriales: la basada en **isohipsas** o contornos y la **red irregular de triángulos** —*TIN*, *triangulated irregular network*— y dos estructuras raster: las **matrices regulares** —*URG*, *uniform regular grids*— y las **matrices jerárquicas** —*quadrees*—:

1. **Estructuras vectoriales**, basadas en entidades/objetos
 1. **Contornos**: polilíneas de altitud constante
 2. **TIN**: red de triángulos irregulares adosados
2. **Estructuras raster**, basadas en localizaciones
 1. **Matrices regulares**: malla de celda cuadrada
 2. **Quadrees**: matrices imbricadas en una estructura jerárquica

Modelo vectorial: contornos.

La estructura básica de un modelo de contornos es la **polilínea** definida como un vector de n pares de coordenadas (x, y) que describe la trayectoria de las curvas de nivel o isohipsas. El número de elementos de cada vector es variable; la reducción de éste a un único elemento, $n=1$, permite incorporar **elementos puntuales** (cotas) sin introducir incoherencias estructurales. Una curva de nivel concreta queda definida, por tanto, mediante un vector ordenado de puntos que se sitúan sobre ella a intervalos adecuados —no necesariamente iguales— para garantizar la exactitud necesaria del modelo. La localización espacial de cada elemento es explícita, conservando los valores individuales de coordenadas. En el caso más sencillo, el MDE está constituido por el conjunto de las curvas de nivel que pasan por la zona representada, separadas generalmente por intervalos constantes de altitud, más un conjunto de puntos acotados que definen lugares singulares —cimas, fondos de dolinas, collados, etc.—.



El modelo de contornos es el utilizado habitualmente en los mapas impresos. Es una estructura de datos poco útil para el manejo por medios informáticos pero la escasez de otras fuentes de datos obliga a que todos los SIG tengan herramientas para incorporarlos y usarlos, normalmente mediante transformaciones a otras estructuras (TIN o matrices)

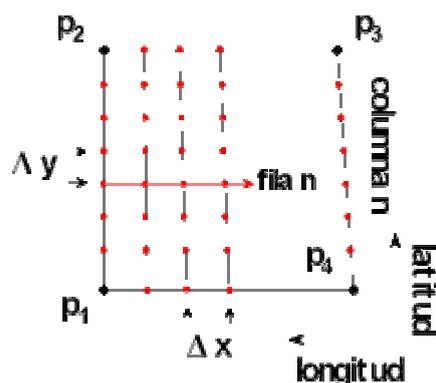
Modelo vectorial: redes de triángulos irregulares (TIN)

Esta estructura de datos se compone de un conjunto de **triángulos irregulares adosados** y que suele identificarse por las siglas de su denominación inglesa: *triangulated irregular network*, *TIN* (Peucker *et al.*, 1978). Los triángulos se construyen ajustando un plano a tres puntos cercanos no colineales, y se adosan sobre el terreno formando un mosaico que puede adaptarse a la superficie con diferente grado de detalle, en función de la complejidad del relieve. Se trata de una estructura en la que el terreno queda representado por el conjunto de superficies planas que se ajustan a un conjunto previo de puntos.

El método de triangulación más utilizado se denomina triangulación de Delaunay y se explica con mayor detenimiento en la parte final de este capítulo.

Modelo raster: matrices regulares

La estructura matricial tiene antecedentes relativamente remotos: Chapman (1952) propone ya métodos de análisis topográfico basados en matrices regulares. Esta estructura es el resultado de superponer una retícula sobre el terreno y extraer la altitud media de cada celda. La retícula adopta normalmente la forma de una **red regular de malla cuadrada**. En esta estructura, la localización espacial de cada dato está determinada de forma implícita por su situación en la matriz, una vez definidos el origen y el valor del intervalo entre filas y columnas.



La matriz regular es la estructura más utilizada para construir los MDE. La razón es que se trata de una estructura de fácil manejo informático y simple de representar mediante estructuras lógicas como matrices de dos dimensiones.

Modelo raster: matrices jerárquicas —*quadrees*—

El mayor interés de las matrices jerárquicas está en que permiten solucionar el principal problema de las matrices regulares: su resolución espacial constante. En este tipo de matrices los elementos pueden ser, bien datos elementales, como en las matrices regulares, o bien, a su vez, matrices de 2×2 con una distancia entre filas y columnas mitad de la del nivel anterior. La estructura final es un árbol jerárquico de matrices elementales de 2×2 con una profundidad en principio arbitraria y cuya resolución espacial se duplica en cada nivel. Esta estructura, denominada originalmente *quadtree*, se ha utilizado ocasionalmente en el tratamiento de variables nominales (Samet *et al.*, 1984) con el fin de reducir el tamaño de almacenamiento.

Este tipo de estructura no ha sido apenas desarrollada ni utilizada, según puede deducirse de la bibliografía, aunque tiene una relación directa con el método fotogramétrico de muestreo progresivo (Makarovic, 1973). En el tratamiento de los MDE, los trabajos pioneros parecen corresponder a Ebner y Reinhardt (1984, 1988), que utilizan un modelo mixto de matrices jerárquicas y estructuras *TIN*.

Otras estructuras

Se han descrito muchas variantes y alternativas para el archivo y tratamiento de los MDE. La representación mediante **perfiles** suele citarse como una posibilidad y algunos autores la evalúan positivamente debido a algunas ventajas teóricas (Yoeli, 1983:21); en la bibliografía actual, sin embargo, no se encuentran ejemplos que la utilicen. La codificación de contornos mediante **ecuaciones polinómicas** fue propuesta en los primeros trabajos sobre modelos digitales (Miller y Laflamme, 1958:437); más recientemente, Walton (1989) propone un método similar basado en una secuencia de **segmentos de Bézier** con el fin de reducir el tamaño de los ficheros vectoriales. Otros, como los polígonos irregulares adosados (Moore *et al.*, 1988) o redes regulares hexagonales (Roessel, 1988), aducen en cada caso ventajas para aplicaciones concretas pero su uso no se ha generalizado hasta el momento.

¿Qué estructura debe usarse?

Para tomar la decisión de la estructura de datos idónea es necesario tener en cuenta algunas de sus implicaciones; entre ellas están las siguientes:

- adoptar una estructura de datos concreta supone decidir el método de construcción del modelo e, indirectamente, sobre qué tipo de información

va a ser representada y cuál descartada

- implica decidirse por un esquema concreto de almacenamiento y gestión informática de los datos, con sus ventajas e inconvenientes
- implica la necesidad de traducir los algoritmos a formas concretas compatibles con la estructura de datos elegida
- supone aceptar las limitaciones de las aplicaciones informáticas para gestionar la información en el formato elegido

Es decir, la elección de la estructura de datos es importante porque condiciona el futuro manejo de la información

Algunos autores han opinado que las diferencias teóricas son reducidas si la resolución es similar (Berry, 1988) ya que se trata en todos los casos de una distribución de puntos acotados. Sin embargo, esto es fijar la atención sólo en los elementos primarios del modelo cuando la diferencia fundamental estriba, lógicamente, en la forma de estructurar los datos, en la complejidad de la referenciación espacial interna o topología aplicada a la información y en los procesos de tratamiento que estas circunstancias permiten o exigen.

Los sistemas de información geográfica y algunas aplicaciones dedicadas expresamente al tratamiento de los MDT usan, en la práctica, sólo dos de las alternativas anteriores: las **matrices regulares** y los **TIN**.

El papel del modelo de contornos ha quedado reducido a ser una etapa intermedia en la captura de información: la de **digitalización del mapa topográfico**. El diseño de algoritmos para el manejo posterior de la información se ha mostrado tan dificultoso que no se considera una alternativa viable para el tratamiento de los datos topográficos (Mark, 1979:34).

Los *quadrees* han despertado un mayor interés y existen aplicaciones comerciales que los usan como estructura básica (por ejemplo, SPANS); sin embargo, a pesar de tratarse de una estructura teóricamente adecuada para el tratamiento de la topografía, también presentan serios problemas en el diseño de algoritmos.

Puede resumirse que, entre las dos alternativas restantes y para realizar operaciones de **análisis espacial**, en los SIG se ha optado por usar matrices regulares en el tratamiento de los MDT. El motivo de esta decisión ha sido que los TIN, aunque capaces de representar con fidelidad la topografía por su capacidad de adaptarse al terreno con un nivel de resolución variable, no permiten efectuar cálculos ni desarrollar modelos con la facilidad y rapidez de las matrices. Éstas se han mostrado como estructuras mucho más simples conceptualmente y más adaptadas al tratamiento informático.

La alternativa adoptada en algunos SIG ha sido usar las matrices regulares divididas en elementos denominados *tiles* —"losas" o "baldosas"— cuyo conjunto cubre el territorio. La división en *tiles* permite una gestión más eficaz de la información ya que no exige el manejo global de una matriz de grandes dimensiones. La resolución de las unidades elementales sigue siendo, sin embargo, fija, y la gestión se complica algo más ya que es necesario mantener y manejar una estructura más compleja que la simple matriz regular.

La construcción del MDE

La **captura de la información** hipsométrica constituye el paso inicial en el proceso de construcción del MDE, e incluye la fase de **transformación de la realidad geográfica a la**

estructura digital de datos. Se trata de una fase de gran trascendencia porque la calidad de los datos es el principal factor limitante para los tratamientos que se realicen posteriormente. Tras obtener los datos, éstos deben ser estructurados para formar el MDE de alguna de las formas presentadas en el apartado anterior.

Captura de datos

Los métodos básicos para la conseguir los datos de altitudes pueden dividirse en dos grupos: **directos** —*primary data*— cuando las medidas se realizan directamente sobre el terreno real, e **indirectos** —*secondary data*— cuando se utilizan documentos analógicos o digitales elaborados previamente. La jerarquía de los métodos más usuales es la siguiente:

1. **Métodos directos:** medida directa de la altitud sobre el terreno (fuentes primarias)
 1. **Altimetría:** altímetros radar o laser transportados por plataformas aéreas o satélites
 2. **GPS:** *global positioning system*, sistema de localización por triangulación
 3. **Levantamiento topográfico:** estaciones topográficas con salida digital
2. **Métodos indirectos:** medida estimada a partir de documentos previos (fuentes secundarias)
 1. **Restitución** a partir de pares de imágenes
 1. **Estereo-imágenes digitales:** imágenes tomadas por satélites
 2. **Estereo-imágenes analógicas:** imágenes fotográficas convencionales
 3. **Interferometría radar:** imágenes de interferencia de sensores radar
 2. **Digitalización** de mapas topográficos
 1. **Automática:** mediante escáner y vectorización
 2. **Manual:** mediante tablero digitalizador

Métodos directos: altímetros, GPS y estaciones topográficas

Altimetros radar y láser

Algunos satélites han incorporado **altímetros** entre sus instrumentos, con lo que se hace posible, al menos teóricamente, el registro directo de los datos de altitud en formato digital. Estos altímetros son aparatos extremadamente precisos: el transportado por el satélite *ERS-1* —*European Remote-Sensing Satellite*— es un radar de 13.8 GHz con un error nominal de apenas unos centímetros (Bruzzi y Wooding, 1990:13).

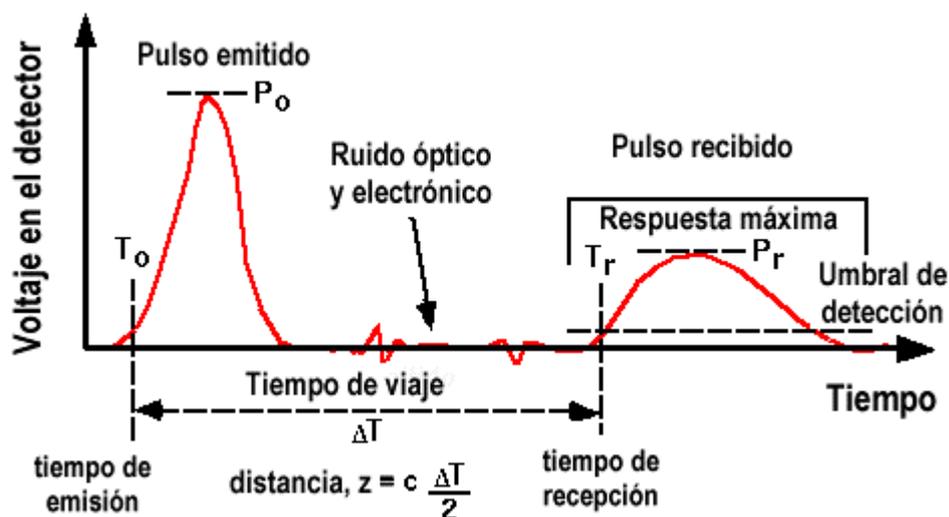
Los **altímetros radar** se sustentan en principios físicos esencialmente diferentes de la teledetección convencional. Los sensores ópticos reciben la fracción de luz reflejada por el suelo ante la radiación electromagnética del Sol; son sensores pasivos (no emiten energía) y la energía que reciben depende del albedo del suelo para el rango de longitudes de onda con el que se trabaja.

El radar es activo: emite una radiación desde una antena emisora y, además, esta radiación es coherente, es decir, se trata de una luz de longitud de onda pura en la que crestas y valles de la onda electromagnética son absolutamente regulares. Debido a esta característica, el sensor puede distinguir el punto exacto en el que se encuentra la oscilación de la onda reflejada: **la fase**.

La medida de la fase es útil debido a la elevada frecuencia de trabajo: si un radar transportado por un satélite funciona a 6 GHz, la señal recorrerá sólo 5 cm en el tiempo de una oscilación completa. Si el suelo está exactamente a 800 km de distancia, los 1600 km de ida y vuelta corresponden a un número entero de ciclos y la onda recibida llegará en fase con la onda de salida. Si la distancia es de 800 km y 1 cm, la onda llegará con un desfase del 40%: un 20% corresponde al cm del viaje de ida y el otro 20% al del viaje de vuelta. Este desfase se detecta en equipo receptor. La medida de la fase permite, por tanto, estimar las distancias con precisiones centimétricas.

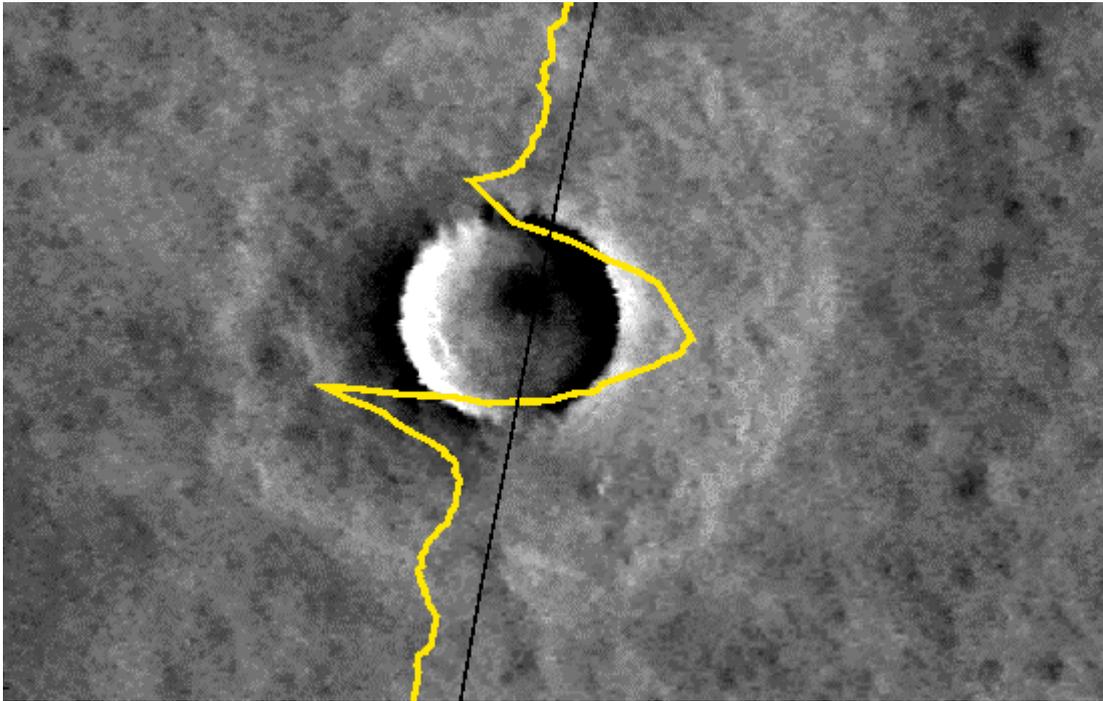
El problema con el que se encuentra esta metodología es que la unidad mínima de medida puede ser un área relativamente amplia de suelo; si este área es heterogénea generará reflexiones múltiples que se combinarán de forma que la fase resultante parece aleatoria e inutilizable. Por estos motivos el uso directo de altímetros ha tenido especial éxito sobre superficies homogéneas como en la medida de la topografía de la superficie marina y en el seguimiento de los hielos polares (Francis *et al.*, 1991:42-43). Veremos posteriormente como esta información ya puede ser aprovechada mediante técnicas algo más sofisticadas.

Los **altímetros laser** utilizados para el levantamiento topográfico están basados en el mismo principio: desde el satélite o avión se lanza un breve pulso hacia la superficie (apenas unos ns); una fracción de la energía es reflejada por la superficie y devuelta hacia una antena receptora. El tiempo transcurrido entre la emisión y la recepción del eco depende de la distancia entre la antena y el suelo.



En la actualidad están en funcionamiento varias misiones basadas en altímetros laser (no necesariamente terrestres). Merece la pena destacar el [Mars Orbiter Laser Altimeter](#) (MOLA) transportado en la misión *Mars Global Surveyor*. Lanzado en noviembre de 1996 y llegado a la

órbita de Marte en septiembre de 1997, está específicamente destinado para la cartografía topográfica de Marte. Aún en la primera fase de la misión, el MOLA ya ha comenzado a enviar resultados. Uno de ellos se muestra en la figura siguiente:



Representación de un perfil altimétrico tomado sobre un cráter de 24 km de diámetro y 786 m de profundidad.

Merece destacarse que la cobertura topográfica más completa disponible adquirida con estos métodos es la de Venus, donde hace ya algunos años, la [misión Magallanes](#) levantó el mapa del planeta prácticamente completo. La nave Magallanes fue lanzada el 4 de mayo de 1989 y llegó a la órbita de Venus unos 15 meses más tarde, el 10 de agosto de 1990. Durante la misión (1990-94) se cartografió el 98% del planeta con una resolución de unos 100 m. Combinando imágenes tomadas en diferentes órbitas del satélite fue posible también crear simulaciones tridimensionales de la superficie de Venus.

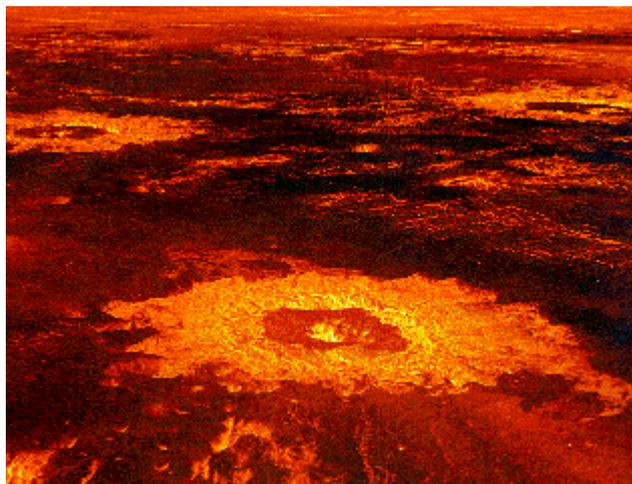


Imagen creada a partir de los datos de la misión Magellan sobre Venus, combinando datos del radar de apertura sintética (SAR) y altímetro. Se muestra a apariencia de 3 cráteres en Lavinia Planitia. El cráter Howe (en primer plano) tiene 37.3 km de diámetro. Los colores se han simulado a partir de datos tomados por las misiones soviéticas Venera 13 y 14.



Altímetro láser transportado por avión.

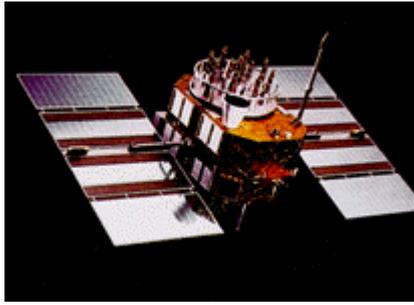
GPS, sistema global de localización

Otro método de cierto interés es el denominado **sistema global de localización**, conocido como GPS —*global positioning system*—. Este método utiliza un conjunto o constelación de satélites de referencia y, mediante métodos de triangulación, permiten obtener valores de las tres coordenadas espaciales para un lugar concreto de la superficie terrestre.

La idea de poder ubicarse en un lugar de la superficie terrestre con gran precisión y en ausencia de referencias locales tuvo su origen con fines militares y, en concreto, fue iniciativa del *U.S. Department of Defense*. El método es esencialmente simple: supongamos que logramos medir que un satélite determinado está a 20.000 km de distancia de nosotros; en ausencia de otra información, nuestra situación potencial queda limitada a la superficie de una gran esfera de ese diámetro. Si medimos la distancia a otro segundo satélite, nuestra potencial ubicación se limita ahora a la intersección entre las dos superficies esféricas: una circunferencia. Si interviene una tercera esfera nuestras posibles posiciones quedan limitadas a 2 puntos y uno de ellos será absurdo por estar lejos de la superficie de la Tierra.

El proceso anterior es simple pero exige determinar la distancia a un mínimo de tres satélites simultáneamente con gran precisión. La opción utilizada es que los satélites emitan pulsos de radiofrecuencia en ciertos momentos y un receptor calcule la distancia en función del momento de recepción. este método exige que **el reloj del receptor esté sincronizado con el reloj del satélite**, concepto básico del sistema global de localización. En Herring (1996) puede encontrarse un artículo de síntesis sobre GPS en español.

La constelación GPS está formada actualmente por un total de 21 satélites primarios y 3 de reserva orbitando a 20.000 km de altitud con un periodo de 12 horas. La precisión potencial es muy elevada, del orden de centímetros, ya que los atélites van provistos de relojes atómicos y se han depurado notablemente los sistemas de sincronización con las estaciones receptoras. Sin embargo, ya en el origen, las Fuerzas Armadas de EE.UU. introdujeron "ruido" en las señales de los satélites de forma que las personas no autorizadas no pudieran ubicarse con demasiada precisión sobre el terreno y aprovechar el sistema para fines militares propios. El método, denominado "disponibilidad selectiva" (*selective availability*) consiste en alterar los relojes de los satélites de acuerdo con un código específico pseudoaleatorio. El objetivo es que un usuario del SPS (*Standard Positioning Service*) pueda ubicarse con unos errores de 100 m (horizontal) y 156 m (vertical) el 95% de las veces.



Satélite de la constelación GPS

Los intentos para corregir estos errores se desarrollaron pronto y utilizaron una técnica simple: midiendo desde un punto fijo y de posición conocida es posible comparar las pseudodistancias (medidas) con las distancias verdaderas (que era posible calcular al saber la posición de la estación terrestre). La medida simultánea con varios satélites permite escapar a los efectos de la "disponibilidad selectiva" y es lo que se conoce actualmente como **GPS diferencial**, con el que la estación terrestre transmite por radio los errores en los relojes para que la estación GPS móvil pueda calcular su ubicación con errores de unos cm.

Curiosamente, esta precisión es notablemente superior a la que pueden obtener los usuarios autorizados para usar el PPS (*Precise Positioning System*) mediante los decodificadores criptográficos que permiten eliminar el ruido inducido en la señal. La precisión que puede esperarse del PPS es de 22 m (horizontal) y 27.7 (vertical).

 [Universidad de Texas: Global Positioning System Overview](#)

En la actualidad, Rusia posee un sistema similar al anterior denominado **GLONASS** (*Global Navigation Satellite System*), también operado por sus Fuerzas Aéreas. El sistema consta de 24 satélites en 3 planos orbitales con una diferencia angular entre planos de 120° y 120°. La constelación está a 19100 km de altitud y completan una órbita a la Tierra cada 11 h y 15 min.

Glonass genera dos tipos de señales o servicios: el SP (*Standard Precision navigation signal*) y el HP (*High Precision Navigation Signal*). El SP está disponible para uso público y permite una precisión horizontal entre 57-70 m y una vertical de 70 m (ambas con probabilidad de 99.7%). La señal HP sólo está disponible para usuarios autorizados.

 [Centro de Información Glonass de Moscú](#)

 [MIT Lincoln Laboratory](#)

A pesar de su precisión, este método presenta algunas limitaciones para la construcción del MDE, entre las cuales destacan la necesidad de acceder físicamente al lugar de medida y el tiempo relativamente elevado que se precisa para realizar una toma de datos precisa. A estas dos circunstancias deben añadirse algunos problemas secundarios relativos a la necesidad de condiciones favorables para las medidas: acceso directo y simultáneo a un mínimo de cuatro satélites, poca cubierta vegetal sobre la antena receptora, necesidad de una segunda estación de apoyo en funcionamiento simultáneo, etc. Estos problemas han convertido al método GPS en un recurso de apoyo, pero no en un sistema básico de captura de datos para la generación de MDE. Sí es un método válido para la construcción del conjunto de puntos de control necesario para la evaluación del error.

Estaciones topográficas

Finalmente, las estaciones topográficas más avanzadas pueden generar y almacenar los

resultados de sus medidas en formato digital. Algunos sistemas de información geográfica incorporan utilidades que permiten el tratamiento e incorporación de estos datos (ver, por ejemplo, ESRI, 1991). A pesar de su utilidad, el método tiene problemas similares al anterior ya que la recogida de información exige la presencia física sobre el terreno. Asimismo, el tiempo necesario para realizar una toma de datos fiable y completa impide que este método pueda ser la única vía de incorporación de datos en la construcción del MDE.

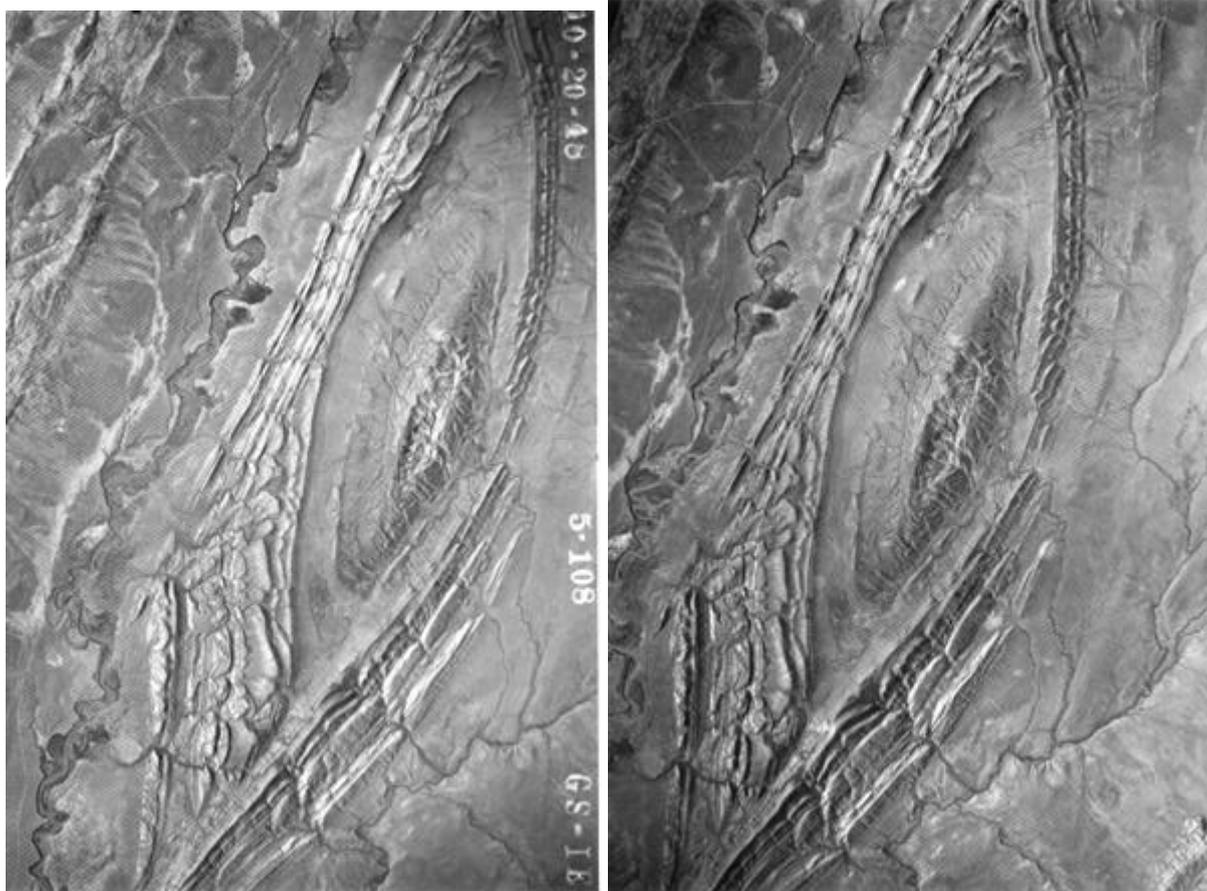
Métodos indirectos: restitución fotogramétrica

Como se ha indicado anteriormente, es bastante más frecuente utilizar métodos indirectos para la generación de los MDE. Las causas son principalmente dos:

- con estos métodos no es preciso acceder físicamente a la zona de estudio, pues utilizan documentos preexistentes, y
- la generación de datos se hace de forma relativamente rápida, cuestión básica cuando el volumen de información es muy elevado.

En las operaciones de restitución fotogramétrica se utiliza como documento básico un conjunto de pares de imágenes de la zona a estudiar, parcialmente solapados y tomados desde puntos de vista diferentes, que forman los denominados **pares estereoscópicos**. Examinando puntos homólogos en los pares estereoscópicos es posible deducir de su paralaje las cotas de referencia necesarias para reconstruir la topografía. Actualmente existen sistemas —restituidores fotogramétricos— automatizados que realizan esta labor grabando directamente los resultados en un formato digital compatible con sistemas de información geográfica. El acceso al terreno es necesario para establecer un conjunto de puntos de apoyo que permitan fijar valores de altitud en una escala absoluta. El método se utiliza también en la elaboración de cartografía analógica y sólo se diferencia en el formato de la salida de los datos. En la cartografía convencional, la salida se realiza sobre un soporte físico estable, mientras que en el otro caso se graba directamente en un soporte informático.

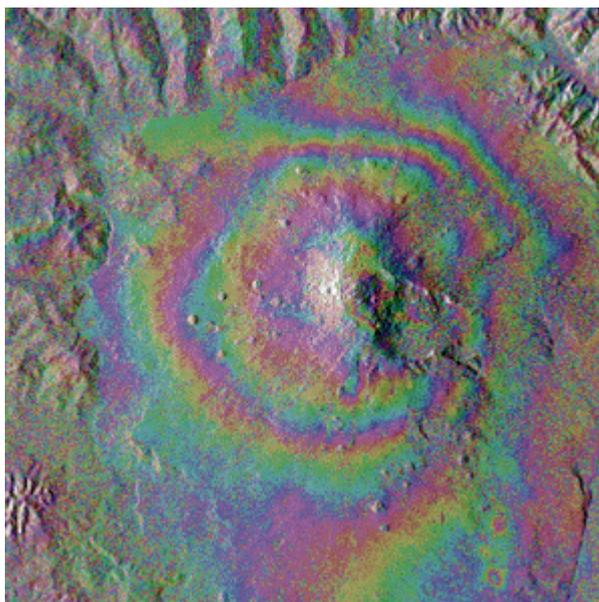
Los pares estereoscópicos han sido hasta hace pocos años exclusivamente fotogramas aéreos, tomados por cámaras de gran formato desde aviones en vuelo a diferentes altitudes. Actualmente, a estas fotografías se han sumado las imágenes digitales tomadas por sensores pancromáticos transportados por satélites. El lanzamiento del satélite SPOT en 1986 ha permitido obtener imágenes estereoscópicas mediante variaciones en el ángulo de vista, programables desde tierra. Los sensores pancromáticos del SPOT permiten una resolución de 10 m, con un rango espectral de visión similar al del ojo humano. Según algunos autores, los MDE resultantes tienen una calidad comparable a los obtenidos mediante los métodos fotogramétricos convencionales (Kaufmann y Haja, 1988). Otros autores, sin embargo, presentan estudios con valores de error más elevados, compatibles solamente con escalas poco detalladas.



Ejemplo de par estereoscópico formado por dos imágenes tomadas desde distintos puntos de vista. La observación mediante un estereoscopio permite observar el relieve de la zona donde ambas imágenes se solapan.

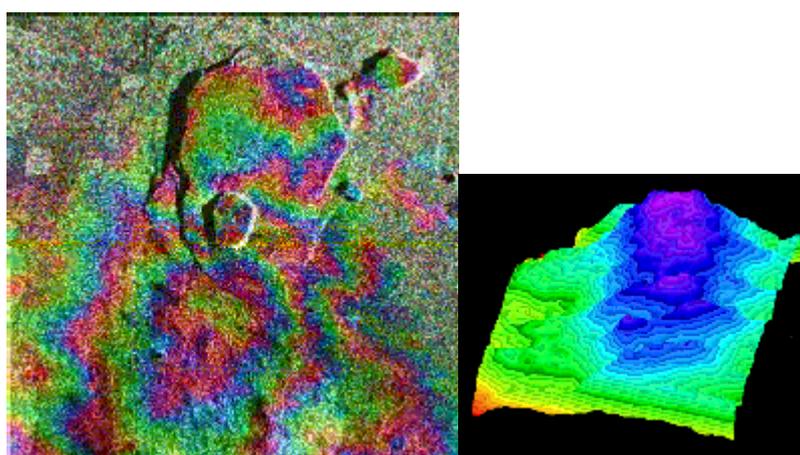
Métodos indirectos: interferometría radar

Ya se han mencionado anteriormente los problemas que existen para estimar la altitud mediante la medida de fase en imágenes radar. Ya en la década de los 80, en el [*Jet propulsion Laboratory*](#) se comenzó a trabajar en una nueva forma de obviar estos problemas. La idea consiste en comparar dos imágenes del mismo lugar tomadas desde la misma posición pero en momentos diferentes. Aunque la fase de cada imagen considerada independientemente era básicamente aleatoria, la **diferencia de fases** exhibía unos patrones de interferencia bastante claros. En principio, si las dos imágenes se toman desde el mismo sitio, no debería haber diferencias de fase entre el mismo pixel de las dos imágenes: las diferencias sólo pueden deberse a cambios ocurridos entre las dos tomas.



El interferograma del Monte Etna (Sicilia), realizado con imágenes del ERS-1 muestra que el monte se hunde según la lava va saliendo; los 4 ciclos de colores indican un hundimiento de 11 cm en el periodo de 13 meses transcurrido entre las dos tomas (años 1992-93). El hundimiento se extiende hasta gran distancia del cráter lo que sirve de indicador de las dimensiones de la cámara subterránea de magma.

De la misma manera, pueden crearse **interferogramas** con dos imágenes tomadas al mismo tiempo pero desde lugares algo diferentes: por ejemplo, a lo largo del paso del satélite o portando dos antenas simultáneamente. En este último caso, las dos antenas tienen vuelos paralelos y están separadas entre si por una distancia determinada. La señal emitida por una de ellas es recibida por las dos antenas simultáneamente. Los diferentes ángulos entre antenas y punto de medida en el suelo inducen una diferencia de fase que puede interpretarse geoméricamente de forma análoga a la usada con imágenes estereoscópicas. También es posible usar una sola antena y comparar imágenes tomadas en dos órbitas diferentes en cuyo caso la distancia entre tomas deben estar dentro de unos rangos determinados para evitar una total pérdida de coherencia; esta distancia crítica depende de la aplicación y se ha valorado en unos 150-300 m para el ERS-1 para la construcción de modelos digitales de elevaciones (Solaas, 1994)



A la izquierda se muestra la imagen interferométrica del volcal Kiluaea (Hawaii) usada para generar el modelo digital de elevaciones mostrado a continuación. En este caso la interferencia se debe a las diferentes órbitas del satélite.

La interferometría radar es extremadamente precisa: Madsen *et al.* (1993), con datos tomados en verano de 1991 consiguieron resultados con un error cuadrático medio (ECM) de 2.2 m para

zonas planas y superiores a 5 m para zonas de montaña. En el caso de radares transportados por aviones, mediante una combinación entre GPS y los sistemas de navegación inerciales, es posible determinar la posición del avión y el ángulo de toma con unas precisiones de 10-20 m y 0.01-0.02 grados respectivamente. Estos resultados permiten iniciar una generación de mapas topográficos con un precisión en la altitud de 2 m o mejor.

Existen 4 satélites con radar operativos actualmente: Radarsat ([CCRS](#), Canadian Center for Remote Sensing), ERS-1 y ERS-2 ([Agencia Espacial Europea](#)) y JERS-1 ([National Space Development Agency of Japan](#)); aunque ninguno de ellos fue diseñado especialmente para generar interferometrías todos pueden ser usados para ello con mayor o menor facilidad y prestaciones.

Como con otros métodos, los primeros pasos de esta técnica se dieron lejos de la Tierra: fue usada inicialmente por Rogers e Ingalls (1969) para la observación de la superficie de Venus y de la Luna. En la actualidad existe un proyecto específico para realizar una cartografía topográfica de la Tierra mediante radares instalados en la lanzadera espacial en el presente año 1999. Este proyecto, denominado **SRTM** ([Shuttle Radar Topography Mission](#)) dependiente del *Jet Propulsion Laboratory* será transportado por la lanzadera espacial que, mediante dos radares tomando medidas simultáneamente a 233 km de altitud cubrirá el 80% de la superficie terrestre (entre los paralelos 60°N y 56°S) en una misión de 11 días de duración. El producto cumplirá las especificaciones ITDH-2 (*Interferometric Terrain Height Data-2*) lo que implica un pixel de 30x30 m con una precisión absoluta altitudinal mínima de unos 12 m. El primer radar se encontrará en el fuselaje de la lanzadera y para llevar a cabo la misión, se desplegará un mástil de 60 m de longitud en cuyo extremo se porta el segundo radar. La misión tiene previsto su lanzamiento el 16 de septiembre de 1999.

Métodos indirectos: la digitalización automática

Los métodos fotogramétricos son utilizados generalmente por organismos de la administración o por empresas especializadas. El aún elevado coste de los aparatos necesarios para la restitución hace difícil que pequeños equipos puedan abordar la construcción de los MDE por esta vía. La opción alternativa es la digitalización de los mapas topográficos preexistentes, lo cual puede realizarse de forma manual, mediante un tablero digitalizador, o automáticamente mediante un escáner.

La **digitalización automática** ha sido una línea de investigación en constante desarrollo en los últimos años debido a que los métodos de digitalización manual son lentos y proporcionalmente muy costosos dentro del total de procesos. Para digitalizar de forma automática, el documento original se analiza mediante un escáner, dispositivo que recorre la superficie del mapa con un sensor óptico y la traduce a un cierto número de niveles de gris o componentes de color. El resultado final es una de valores de reflectancia. El tamaño de la celda o pixel se establece de forma que sea capaz de recoger todas las estructuras presentes en el mapa manteniendo de unas dimensiones de los ficheros tratables por los medios informáticos disponibles.

De forma general, la digitalización de las curvas de nivel se realiza a en un proceso formado por varias fases:

- toma de la imagen mediante el escáner, con un tamaño de pixel suficiente para recoger las líneas y discriminar la mínima separación entre ellas. En Peled y Fradkin (1994:249) se sugiere una resolución de 1000-1200 puntos por pulgada para la digitalización de originales complejos a escala 1:50.000. La imagen puede ser binaria pero también es posible tomarla en color para poder separar elementos de diferentes colores

- filtrado de la misma para separar las estructuras de interés y descartar el resto. Por ejemplo, si las curvas de nivel están en cierto color, se indicarán unas tolerancias para el mismo y se borrarán todos los píxeles que queden fuera del rango.
- vectorización de la imagen contrastada, de forma que los píxeles adyacentes se estructuran en líneas, generándose la imagen vectorial más compacta y utilizable. Su similitud con el mapa original es función de la calidad gráfica y complejidad de éste pero suele ser un producto poco depurado debido a las dificultades de separar las líneas de nivel de cualquier otra entidad del mapa: retículas, toponimia, cotas, carreteras, red hidrográfica, tramas, etc. Esta fase puede realizarse de forma automática o asistida, donde la toma de decisiones ante situaciones conflictivas la realiza un operador ante la pantalla gráfica
- reconocimiento de los valores de elevación escritos en el mapa, en un proceso de varios pasos:
 - una aplicación OCR —*optical character recognition*— identifica los dígitos del mapa
 - se restituyen los rótulos de elevaciones a partir de los dígitos
 - se asigna la elevación a los contornos a ambos lados del rótulo
 - se unen ambos extremos de contornos y se eliminan los rótulos
 - edición manual, para la corrección de los errores topológicos y la asignación de altitudes a las líneas no etiquetadas

De este conjunto de procesos resulta la versión digital del mapa topográfico original cuyo coste y utilidad son factores variables, en los que la elección de un mapa original de buena calidad es decisiva.

El principal problema de la digitalización mediante escáner es que los originales suelen ser complejos, con mucha más información que la estrictamente deseable, que sería exclusivamente las curvas de nivel y puntos acotados. Las confusiones con el resto de elementos del mapa hace que la fase de revisión y corrección sea frecuentemente muy costosa, hasta el punto de que puede tener un rendimiento menor que la digitalización manual. Este tipo de errores se propaga al documento digital como **errores topológicos** (Peled y Fradkin, 1994:246), el problema más frecuente de la digitalización automática.

Por este motivo, y a pesar del coste cada vez más asumible de los escáneres de gran formato, la digitalización manual es un método usado muy frecuentemente por los equipos investigadores de pequeña entidad.

Métodos indirectos: la digitalización manual

La digitalización manual se realiza con un **tablero digitalizador** sobre el que se coloca el mapa. Las curvas de nivel se siguen manualmente con un cursor de forma que el ordenador recibe a ciertos intervalos, prefijados o decididos por el operador, las coordenadas que definen la trayectoria de la línea.



Los tableros digitalizadores se componen de una densa rejilla que es capaz de localizar el cursor con una precisión elevada (variable según el modelo pero no inferior a los 0.25 mm). Esta precisión es superior a la que un operador es capaz de mantener durante una sesión de trabajo.

Los programas de digitalización actuales tienen capacidad suficiente para permitir trabajos complejos. Existen, sin embargo, otros factores que influyen en la eficacia global del proceso, como los siguientes:

- la capacidad del ordenador utilizado, que debe estar adaptada a la complejidad real de los mapas. Aunque los SIG suelen incluir utilidades de digitalización, su uso obliga habitualmente a inmovilizar un ordenador de gran capacidad para una tarea que no la necesita, lo que supone un derroche de medios poco justificable.
- el uso de programas específicos para la digitalización obliga frecuentemente a usar formatos de almacenamiento que, aunque adaptados al manejo de la información gráfica, no se pueden usar de forma directa por los SIG. Este problema exige que se prevean **métodos de conversión de formatos** que ocasionen una pérdida mínima de información. Debe resaltarse que los procesos de traducción de formatos, sobre los que se suele hacer poco énfasis, pueden suponer un problema no trivial y que exigen una inversión de tiempo a veces muy elevada.

Los trabajos de digitalización son, en la práctica, de calidad muy irregular. Aunque la competencia profesional del operador es el factor básico, existen algunas normas que ayudan a conseguir un trabajo de mayor calidad. Por ejemplo:

- deben usarse mapas en buen estado, ya que la precisión del trabajo depende de la calidad del documento original. Deben evitarse, por tanto, los mapas doblados, deformados o con las líneas poco visibles.
- la referencia espacial del mapa debe registrarse con precisión mediante puntos de control localizados en la periferia del área, con la máxima separación posible en el plano XY .
- debe aprovecharse la capacidad de los SIG actuales, que permiten incluir un conjunto variado de elementos a la hora de generar el MDE: a las curvas de nivel se unen los puntos acotados, la red hidrológica, líneas de ruptura, límites de zona, fragmentos de curvas de apoyo, etc.
- debe valorarse el método de digitalización a utilizar, incremental o punto a punto, en función de la calidad de los resultados.
- no es deseable introducir un número excesivamente elevado de puntos

dentro de las curvas de nivel. En principio, los resultados óptimos en los procesos de generación del MDE se obtendrán cuando la distancia entre puntos a lo largo de una línea sea similar a la distancia entre líneas (Flach y Chidley, 1988).

Este último factor diferencia la digitalización de las curvas de nivel de la de mapas temáticos, donde las fronteras entre categorías debe definirse con la mayor precisión posible.

Elementos que definen la estructura topográfica

Los MDE manejados por los SIG actuales permiten solamente un valor de altitud por lugar geográfico, es decir, no se admiten pliegues ni estructuras estrictamente verticales. Los MDE no son, por tanto, modelos 3D verdaderos, como los usados habitualmente en CAD. Aunque se han diseñado SIG capaces de manejar datos en tres dimensiones reales, aún no se han generalizado y han quedado restringidos a prototipos experimentales, generalmente de aplicación geológica.

Antes de comenzar la operación propiamente dicha de generación del MDE, el fichero resultante de la digitalización debe someterse a una serie de operaciones de estructuración y asignación de atributos que, en resumen, son las siguientes:

- **estructuración topológica**, mediante la cual se crea la base de datos necesaria para mantener y gestionar las relaciones línea-nodo
- **asignación de altitudes** a las líneas y puntos, ya que este atributo puede no transferirse durante la digitalización
- la **generalización**, un proceso de filtrado selectivo mediante el cual se eliminan la información innecesaria o redundante en función del error admisible
- **asignación de códigos** a las estructuras auxiliares que hayan sido digitalizadas como información complementaria

Una vez realizadas estas operaciones es posible abordar la creación del MDE según métodos diversos. La experiencia ha mostrado que la calidad de un MDE puede mejorarse significativamente introduciendo **datos auxiliares** de diversos tipos como, por ejemplo

- curvas de nivel o *contours*, que habitualmente se descomponen en puntos acotados —*mass points*— previa generalización o reducción de la densidad de vértices de la línea
- puntos acotados singulares o *vips* —de *very important points*—, que definen puntos con información relevante; por ejemplo, cumbres de picos o fondos de dolinas. En ocasiones ambos tipos de datos, cumbres y sumideros deben ser diferenciados claramente debido a su diferente tratamiento.
- líneas estructurales, que definen elementos lineales *con valores de altitud* asociados a cada vértice y utilizados para asegurar el ajuste de la superficie topográfica; por ejemplo, la red fluvial o las líneas de cresta —este tipo de líneas han sido utilizadas frecuentemente para la representación del terreno en libros de viajes o mapas excursionistas—.
- líneas de inflexión o rotura —*breaklines*—, utilizadas para definir la

posición de elementos lineales *sin valores de altitud* explícitos que rompen la continuidad de la superficie; por ejemplo, bordes de acantilados. Estas líneas no aportan valores de altitud pero sí intervienen en el comportamiento de los algoritmos de interpolación bloqueando la búsqueda de datos. Una parte de las líneas de inflexión pueden ser extraídas de mapas geomorfológicos.

- zonas de altitud constante, definidas por polígonos que encierran una superficie plana de altitud única; por ejemplo, la superficie de un lago
- zonas de recorte, que define los límites externos del MDE por fuera de los cuales cualquier dato existente se ignora
- zonas vacías, donde las circunstancias impiden la obtención y asignación de cotas; por ejemplo, zonas inundadas, cubiertas de nieve o donde los datos de origen son de calidad dudosa

El conjunto de datos formado por las estructuras mencionadas puede servir para generar un MDE que aproveche adecuada y eficazmente la información topográfica disponible. Debe destacarse que la totalidad de las estructuras anteriores son vectoriales: puntos, líneas y polígonos.

Métodos de construcción del MDE

La construcción de un MDE entendido como matriz regular a partir de la información vectorial mencionada anteriormente es básicamente un problema de **interpolación**: se definen las localizaciones de los puntos problema —intersecciones de filas y columnas— y se estima la altitud de cada uno de ellos en función de los datos del entorno existentes en el modelo vectorial. Los algoritmos de interpolación deben tener en cuenta, asimismo, las diversas estructuras auxiliares para introducir variantes en los mecanismos de cálculo.

El planteamiento global del problema podría ser, por tanto, como sigue:

- dado un conjunto de puntos con coordenadas (x, y, z) , distribuidos irregularmente, generar un nuevo conjunto de puntos localizados en los nodos de una red regular de forma que la superficie interpolada sea una representación de la original con una pérdida mínima de información.

En el caso de la creación de una estructura TIN, en lugar de interpolar, se realiza una selección de puntos descartando los que no aporten una información relevante para la descripción de la altitud.

Sin embargo, en un modelo vectorial como el que forma el mapa topográfico digitalizado, acceder directamente a un dato por su posición espacial no es sencillo. Por este motivo, la localización de los datos incluidos en el entorno del punto problema es una operación clave, que debe ser bien diseñada dado el enorme número de datos que suele componer el mapa original. Las estrategias de búsqueda son un problema informático más que geográfico por lo que no entraremos en ellas en este texto.

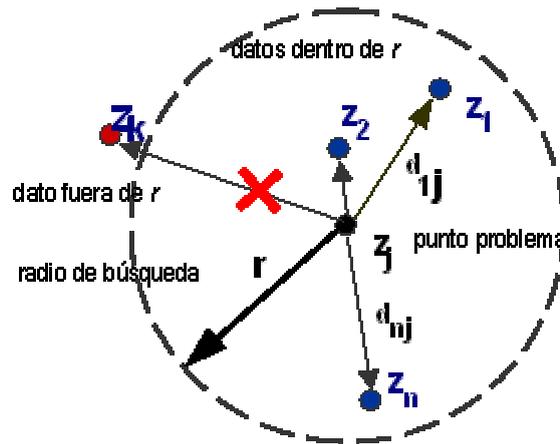
Interpolación en función de la distancia

El valor del punto problema se estima asignando pesos a los datos del entorno **en función inversa de la distancia** que los separa del punto problema —*inverse distance weighting, IDW*—. Se establece, por tanto, que los puntos más cercanos tienen un peso mayor en el cálculo, aunque la relación no tiene porque ser lineal.

El método general

La fórmula general de la interpolación en función inversa de la distancia es:

$$\hat{z}_j = \sum_{i=1}^n k_{ij} \cdot z_i$$
 donde \hat{z}_j es el valor estimado para el punto j ; n es el número de puntos usados en la interpolación; z_i el valor en el punto i -ésimo y k_{ij} el peso asociado al dato i en el cálculo del nodo j . Los pesos k varían entre 0 y 1 para cada dato y la suma total de ellos es la unidad.



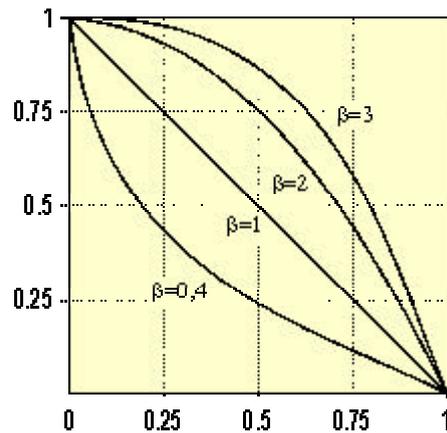
Localización de datos para la interpolación mediante un radio de búsqueda. Los puntos más distantes son descartados. Otras variantes del método incluyen la búsqueda por cuadrantes, la definición de un número de datos mínimo y diversas formas de ponderar las distancias.

Las diferencias entre los diversos métodos estriban en la forma de calcular los pesos de cada dato. Los métodos de distancia inversa calculan la distancia euclidiana entre cada dato y el punto problema, d_{ij} —ver figura anterior—; al establecer una función de proporcionalidad entre el peso y la distancia, la fórmula general queda como sigue:

$$\hat{z}_j = \frac{\sum_i \frac{z_i}{d_{ij}^\beta}}{\sum_i \frac{1}{d_{ij}^\beta}}$$

donde β es un exponente de ponderación. Este parámetro controla la forma en la que el peso disminuye con la distancia. En el caso de valer 0, el valor estimado será la media aritmética de los datos; en el caso de valer 1, el peso disminuye linealmente con la distancia; para valores superiores, la influencia de los puntos cercanos se hace mucho mayor que la de los lejanos. Por ejemplo, cuando $\beta=2$, la interpolación se realiza en función inversa del cuadrado de la distancia.

Cuando se aplican estos métodos aparece frecuentemente un problema derivado del desigual reparto espacial de los puntos a lo largo de las líneas y entre ellas, tal como se ha señalado en el apartado anterior. Se trata del "aterrazamiento" del terreno debido a la ausencia de datos entre curvas de nivel sucesivas. Cuando la interpolación se realiza mediante funciones de ajuste más complejas —polinomios de grado alto—, pueden aparecer artefactos aún más molestos, con fuertes fluctuaciones en la altitud.



Ejemplos de las curvas generadas para varios valores de β donde pueden observarse los patrones de cambio de los pesos (ordenadas) en función de las distancias (unidades arbitrarias, en abscisas)

Esta familia de métodos permite la generación del MDE de una forma rápida y simple. Sin embargo, se trata esencialmente de una media ponderada y, por tanto, el resultado se encuentra siempre incluido dentro del rango de variación de los datos. Por este motivo, el correcto tratamiento de las formas cóncavas y convexas depende estrechamente de la distribución de los puntos originales y la presencia de datos auxiliares se hace muy conveniente.

Por otra parte, la selección del valor de β se hace frecuentemente de forma arbitraria, y al cambiar los valores de β se generan superficies topográficas muy diferentes. La selección objetiva de los pesos es el punto fuerte del método que se muestra en el apartado siguiente.

La hipótesis de la variable regionalizada

La última alternativa para la interpolación que trataremos aquí se denomina habitualmente *kriging*, término construido a partir del nombre de D.G. Krige, uno de los pioneros del método —en español se utilizan también las expresiones "krigeado" y "krigeage"—.

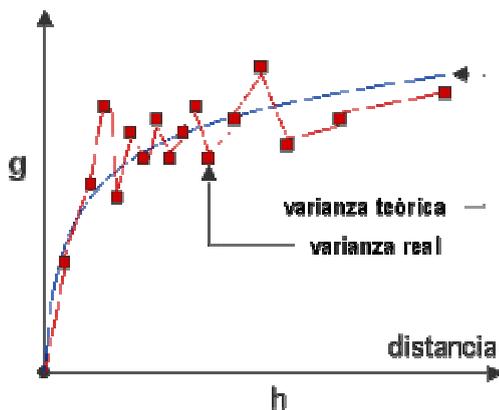
El *kriging* es un método de interpolación con la misma expresión general que los mencionados en el apartado anterior. La diferencia básica es que asume que la altitud puede definirse como una **variable regionalizada**. Esta hipótesis supone que la variación espacial de la variable a representar puede ser explicada al menos parcialmente mediante funciones de correlación espacial: la variación espacial de los valores de z puede deducirse de los valores circundantes de acuerdo con unas funciones homogéneas en toda el área.

Con este método, las funciones pueden deducirse analizando la **correlación espacial** entre los datos en función de la distancia entre ellos. En efecto, en el caso de los MDE, es razonable suponer que el valor de altitud en un punto está relacionado de alguna manera con el valor de los puntos vecinos, distribuidos a distancias variables. Puede suponerse, asimismo, que la "influencia" de los puntos más lejanos es menor que la de los más próximos. El *kriging* estima esta dependencia mediante un estadístico: la **semivarianza** entre datos separados por distancias diferentes.

La semivarianza adquiere valores diferentes en función de la distancia entre los datos (a más separación, menos correlación entre los valores de altitud). La función que relaciona la semivarianza γ con la distancia h se denomina **semivariograma** y muestra, por tanto, la variación de la correlación entre los datos en función de la distancia. La expresión más usual para representarla es (Oliver y Webster, 1990:315):

$$\hat{\gamma}(h) = \frac{1}{2 \cdot n} \sum_{i=1}^n (z_i - z_{i+h})^2$$

donde n es el número de pares de valores separados por una distancia h . Cambiando progresivamente este último valor, es posible elaborar el **semivariograma empírico** para la zona concreta que trabajemos. Normalmente, este semivariograma no se utiliza directamente sino que se examina un conjunto de funciones teóricas y se usa la de mejor ajuste con los datos reales. Finalmente, los pesos k correspondientes a cada distancia entre datos se estiman a partir de los valores del semivariograma por métodos de cálculo matricial (ver Royle *et al.*, 1981 para un desarrollo más detallado).



Ejemplo de semivariograma donde la variancia real se ajusta a una distribución teórica; ésta es la que se aplica para la estimación de los pesos en la interpolación mediante *kriging*

El *kriging* permite obtener dos valores de interés. En primer lugar, el semivariograma nos permite conocer el valor de distancia h para el cual los datos pueden considerarse ya independientes entre sí. Este valor de distancia pone un límite superior al área del entorno alrededor del punto problema, ya que la inclusión de puntos situados a distancias mayores no aportaría información de interés. El otro parámetro estimado es el valor del error esperado para cada punto, que es función de los valores de semivarianza observados para los datos y que puede ser calculado sin dificultad.

Aunque el *kriging* es un método de estimación óptimo desde el punto de vista estadístico, presenta algunas dificultades a la hora de ser utilizado como método de construcción de los MDE. La más importante es que la teoría exige que el semivariograma sea de validez general para toda el área del MDE; esto quiere decir que la interdependencia entre los datos debe ser función exclusivamente de la distancia entre ellos —de su posición relativa— y no de su localización espacial absoluta. Los patrones de variación de z deben ser, por tanto, homogéneos. Esta condición no suele cumplirse y, además, no permite el tratamiento de discontinuidades topográficas que supongan cambios bruscos, como rupturas de pendiente.

Las variables en las que está indicado el uso del *kriging* son aquellas cuya distribución tiene una fuerte componente estocástica o que son resultado de un conjunto de procesos sometido a la interacción de múltiples variables concurrentes. Por ejemplo, en la distribución de un contaminante atmosférico influyen complejas causas meteorológicas a muy diferentes escalas y sometidas a fenómenos de turbulencia. A pesar de que la modelización de estos procesos puede ser muy compleja, los valores puntuales pueden ser descritos satisfactoriamente mediante el *kriging* dado que este tipo de variables cumple las condiciones anteriormente mencionadas.

Los métodos basados en triangulaciones

Los métodos descritos anteriormente tienen como resultado un MDE raster. Algunos de ellos permiten incorporar parte de las estructuras auxiliares como las zonas de recorte o, incluso, líneas de ruptura. Sin embargo, todos los métodos basados en pesos son básicamente filtros ponderados y responden mejor ante variables con distribuciones suaves. En el caso de la topografía, cuando se está tratando con una zona con relieve fuerte, con discontinuidades e, incluso, cuando se desean introducir en el modelo estructuras artificiales de diseño lineal, los métodos descritos no son satisfactorios.

La red irregular de triángulos

La solución aparentemente más adecuada hasta el presente ha sido el tratamiento del relieve mediante una estructura vectorial especial: la **red irregular de triángulos** o TIN —siglas de *triangulated irregular network*—, propuesta inicialmente por Poiker *et al.* (1978).

El modelo TIN tiene varias cualidades de gran interés para los MDE:

- no presupone ni exige la continuidad estadística de la superficie a representar
- puede generarse incorporando una amplia variedad de estructuras auxiliares, especialmente las líneas estructurales y de inflexión, que son incorporadas al modelo como lados de triángulos
- se adapta a la complejidad local del terreno, creando redes localmente más densas en función de la anfractuosidad del relieve
- respeta los valores de los datos, que son usados como vértices y mantienen su altitud exacta

Como ya hemos visto, un TIN es un conjunto de triángulos adosados que recubren el terreno de forma completa. La creación de un TIN se compone de dos partes diferentes: la triangulación propiamente dicha y la selección de los puntos que deben usarse como vértices.

La triangulación ha sido estudiada desde hace más de un siglo y, actualmente, es un tópico bien conocido utilizado en numerosas aplicaciones como robótica, análisis de elementos finitos, visión artificial y síntesis de imágenes (Cignoni *et al.*, 1994:2). En el caso de los SIG el método más habitual es el conocido como **triangulación de Delaunay**, TD.

La TD se entiende mejor partiendo de un concepto complementario, la **teselación de Voronoi**. Los argumentos básicos son los siguientes:

- dados 2 puntos P_i y P_j , en un plano T , la perpendicular al segmento P_iP_j en su punto medio divide el plano en dos regiones V_i y V_j ; la región V_i contiene todos y sólo los puntos cuya distancia a P_i es menor que a P_j ; la región V_j contiene el resto.
- el concepto se extiende a múltiples puntos P_n de forma que cada uno de ellos se asocia a una región V_n que contiene todos los puntos del plano más próximos: $\forall p \in V_i \Rightarrow d(p, P_i) < d(p, P_{x \neq i})$
- si el concepto se aplica a un dominio cerrado se genera un conjunto de polígonos convexos que teselan el plano denominados **regiones de**

Voronoi

- finalmente, si se conectan entre sí los puntos que comparten un borde de una región se obtiene una **triangulación de Delaunay**

La TD genera una solución única para un conjunto de puntos dado salvo en algunos casos especiales —por ejemplo, cuando cuatro puntos pertenecen a una circunferencia— donde pueden existir triangulaciones alternativas válidas. En cualquier caso, las propiedades del método garantizan una forma idónea de los triángulos. Las dos propiedades principales de la TD son:

- un punto P perteneciente a la red es siempre un vértice de un triángulo T
- dos triángulos de una red pueden estar conectados por un vértice o un lado común: no existe superposición
- dado un triángulo $T(P_i, P_j, P_k)$ no existe otro punto de la triangulación interno al círculo que pasa por P_i, P_j y P_k
- dados 4 puntos y el cuadrilátero definido por ellos, la diagonal que lo divide en dos triángulos es aquella que hace máximo el menor de los ángulos internos, es decir, la TD tiende a crear triángulos lo más equiláteros posible

La figura inferior permite ver una triangulación de Delaunay creada sobre la marcha según se van introduciendo puntos sobre el "terreno". Las líneas amarillas forman la triangulación y las líneas azules definen la **teselación de Voronoi o de Thiessen**, un concepto complementario del anterior y usado en los sistemas de información geográfica en algunas operaciones con estructuras vectoriales.

Introduzca puntos con el ratón sobre la superficie gris. Con el botón reset se borrarán todos los puntos y se podrá empezar de nuevo. Si desactiva la opción *triangulations* sólo se mostrará la teselación de Voronoi.

El problema de la triangulación es la selección de puntos que deben usarse como vértices. La forma más trivial de construir un TIN es usar todos los vértices y nodos de las curvas de nivel, así como los puntos singulares como vértices de triángulos. Este modelo masivo no es deseable, en primer lugar, por la enorme cantidad de elementos que sería necesario construir y manejar para una zona de cierta extensión. Es muy probable, asimismo, que fuera altamente redundante ya que muchos puntos procedentes de una digitalización rutinaria no aportan una información significativa.

Formalmente, el problema puede plantearse como sigue: dado un conjunto \mathbf{H} de n puntos acotados, elegir un subconjunto mínimo \mathbf{S} de \mathbf{H} constituido por m puntos, a partir del cual pueda reconstruirse \mathbf{H} con el menor error posible. Los dos enfoques básicos son:

- hacer la selección antes de la triangulación de forma que los algoritmos para la TD conozcan previamente el conjunto completo de vértices y realicen posteriormente un modelo masivo, y
- utilizar un algoritmo de triangulación que realice la selección según realiza la construcción de la red.

En el primer caso, la estrategia es suministrar al algoritmo de TD los puntos básicos ya elegidos. Esto puede realizarse mediante la adecuada generalización cartográfica de las curvas de nivel y

aportando una selección de puntos críticos que representan

Los métodos de construcción de TIN más utilizados pueden agruparse en las siguientes clases:

- **inserción incremental** —*multi-pass refinement methods*—, que comienza con una triangulación mínima y a la que se añaden progresiva y selectivamente nuevos puntos como vértices de la red. El proceso continúa hasta conseguir un objetivo concreto, definido en función de un valor umbral de error.
- **reducción selectiva** —*multi-pass decimation methods*— por eliminación de puntos a partir de un modelo masivo mediante criterios de incremento mínimo del error.

Aunque la triangulación de Delaunay es probablemente el método más utilizado en los SIG actualmente, se trata de un método bidimensional puro: los triángulos se construyen con criterios geométricos en el plano de proyección, usando sólo las coordenadas x e y , independientemente del valor de altitud. Por este motivo, la TD se integra entre los métodos que sólo tienen en cuenta la forma de la triangulación: *shape-dependent methods*.

Dado que parece lógico utilizar los valores de z en un proceso cuyo objetivo es precisamente representar un campo de altitudes, se han desarrollado métodos para realizar la triangulación de forma que el trazado se define en un contexto 3D. Estos métodos se denominan genéricamente *data-dependent methods*. En estos casos, la validez de los nuevos triángulos no se comprueba mediante el test del círculo inscrito sino comprobando cuál de las opciones reduce el error del ajuste en mayor medida. Este tipo de métodos, basado exclusivamente en la reducción del error tiene el problema de que no garantiza una triangulación global óptima como en el caso de la TD. Asimismo, al no tender a maximizar el ángulo mínimo genera frecuentemente un elevado número de triángulos con lados muy desequilibrados —*sliver triangles*—. En cuanto a la calidad de los resultados, Garland y Heckbert (1995:27) indican que estos métodos son significativamente mejores en el caso de superficies de baja rugosidad —formas geométricas, por ejemplo—, para las que es posible conseguir un error similar con muchos menos triángulos. Se indica, sin embargo, que para la modelización de superficies naturales, mucho más rugosas por lo general, la triangulación de Delaunay ofrece resultados de calidad similar.

Las transformaciones vector-raster

El problema más importante de los modelos basados en triángulos es que generan una estructura más difícil de manejar que la matriz regular, especialmente en procesos de análisis, superposición y combinación temática. Es decir, a pesar de ser el TIN la estructura más adecuada para construir un MDE por la capacidad de representar un relieve complejo con precisión, está sometida a los problemas típicos de las estructuras vectoriales.

Por este motivo, y por la ausencia de un desarrollo suficiente de algoritmos de análisis, lo más usual es crear un MDE según el modelo TIN y, posteriormente, generar un MDE matricial convencional mediante procesos de interpolación específicos. Asimismo, es frecuente realizar una selección de puntos críticos a partir de datos matriciales para "alimentar" una triangulación de Delaunay.

En el primer caso, el proceso de interpolación asigna a cada nodo del MDE matricial un valor de altitud estimado a partir del TIN. Obviamente, la estructura matricial no puede representar puntos singulares ni estructuras lineales como el TIN, por lo que parte de los beneficios de la triangulación se pierden. Sin embargo, el proceso aprovecha parcialmente la capacidad del TIN para integrar discontinuidades en la generación del MDE por lo que es preferible a la generación

directa de la matriz regular a partir de los datos de entrada.

La interpolación puede realizarse por dos métodos básicos:

- **interpolación lineal**, donde la altitud del punto problema se estima directamente a partir la ecuación del plano definido por los tres vértices del triángulo que lo contiene; en el caso de que el punto coincida exactamente con un vértice, se le asigna la altitud de éste. La interpolación lineal, aunque muy rápida, provoca la discontinuidad entre triángulos, con cambios bruscos de orientación y pendiente.
- **interpolación quíntica**, que considera la superficie definida por el TIN como un continuo y, por tanto, suaviza las zonas de los vértices y lados; para esta interpolación se utiliza una ecuación polinómica bivariable de quinto grado. En algunos SIG se ha incorporado la posibilidad de integrar las líneas de rotura de forma que rompan la continuidad realizando una interpolación lineal a lo largo de la línea. Salvo en esta circunstancia, la interpolación conserva la continuidad de la pendiente entre triángulos, generando un MDE sin facetas.

Conservando la continuidad hidrológica

La experiencia en la creación de los MDE ha mostrado algunos problemas recurrentes que ocasionan dificultades en el uso posterior de los modelos. Uno de ellos es el aspecto aterrazado ya mencionado anteriormente, debido a un desequilibrio en la distribución espacial de los datos y el tamaño de muestreo elegido para crear la matriz regular. Otro problema es la aparición de **concavidades o pozos** a lo largo de los fondos de valles, lo que les da un característico aspecto granuloso. Este tipo de artefactos se generan por el uso de funciones de interpolación de grado superior en zonas conflictivas.

La existencia de una concavidad puede responder a una forma real del terreno pero, más frecuentemente, se debe a artefactos derivados de la construcción del MDE. Band (1986:16) los atribuye a los procesos de restitución fotogramétrica e indica que, salvo en zonas afectadas recientemente por procesos kársticos y glaciario, las depresiones pueden considerarse errores en los datos. La presencia de concavidades tiene importancia en el caso de los MDE destinados a la simulación de procesos hidrológicos ya que interrumpen las líneas de flujo.

La corrección de concavidades de los MDE ya existentes es posible mediante algoritmos que simulan la inundación y relleno de los pozos (Jenson y Domingue, 1988:1594). El proceso modifica el modelo de elevaciones original por lo que en algunos casos —zonas kársticas, por ejemplo— debe ser aplicado con precaución y los resultados deben interpretarse teniendo en cuenta el contexto en el que se obtienen.

Lo ideal es, obviamente, la generación del MDE mediante métodos que tengan en cuenta el problema y lo solucionen en origen. Una forma de conseguirlo es introducir la red hidrológica como líneas de rotura en la generación del MDE. De esta forma, la triangulación utilizará estas líneas como lados de los triángulos, asignándoles las altitudes correspondientes y trazando correctamente las líneas de flujo.

Hutchinson (1989:213) propone un método específico para la creación de un MDE libre de pozos que impone restricciones al proceso de interpolación forzando su ajuste a la red hidrológica. El trazado fluvial debe ser introducido como información auxiliar.

El SIG Arc/Info ha integrado este proceso en una de sus rutinas de forma que permite la

eliminación de todas las depresiones que no se identifiquen explícitamente como tales. La localización de las depresiones reales debe ser incorporada como dato auxiliar para que el algoritmo las tenga en cuenta, así como el trazado de la red hidrológica. La eliminación de pozos queda limitada optativamente mediante un valor de tolerancia que limita la profundidad de las depresiones a rellenar.

El mayor problema de este método es que, en su estado actual, no permite la introducción de todas las estructuras auxiliares mencionadas al comienzo de este apartado.

FIN DEL CAPÍTULO