

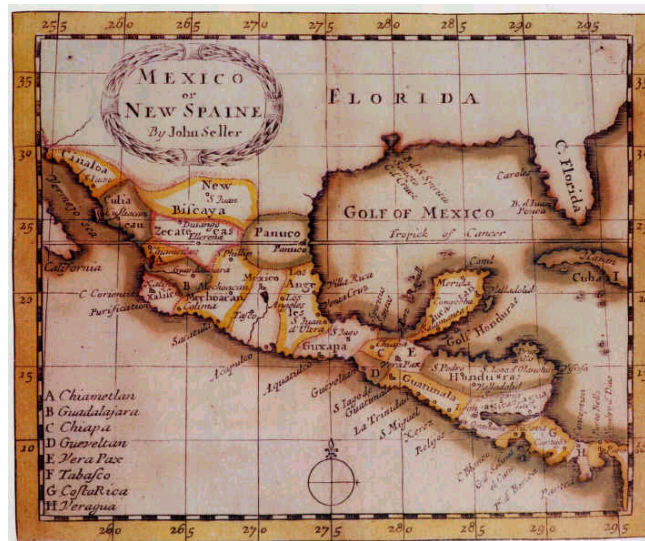
# Introducción

## 2.1. Origen de los mapas

Los símbolos impresos en distintos medios, han sido la forma más utilizada desde tiempos remotos para transmitir ideas, los hombres han usado mapas desde la más remota antigüedad, y probablemente ya los hacían en épocas prehistóricas, es posible que incluso algunos dibujos encontrados en cuevas y refugios, con un significado desconocido hasta el momento, son posiblemente croquis de los territorios donde vivían y cazaban.

Debido a la necesidad que tiene el hombre de conocer la configuración de la Tierra y los accidentes geográficos que en ella existen, surge la necesidad de su representación. Los mapas en papel son un medio fácil para representar nuestro entorno en una forma analógica, que requieren de la intervención humana para su creación e interpretación.

Los mapas han evolucionado desde la antigüedad clásica, Europa medieval, el renacimiento, el siglo diecinueve, el mundo moderno y la revolución producida por las computadoras, los cambios son más bien en su forma que en su construcción, también han cambiado en áreas tales como exploración, guerra, administración del gobierno y la ciencia [Cervera 1999<sup>2</sup>].



John Seller, "MEXICO OR NEW SPAINE". 1679

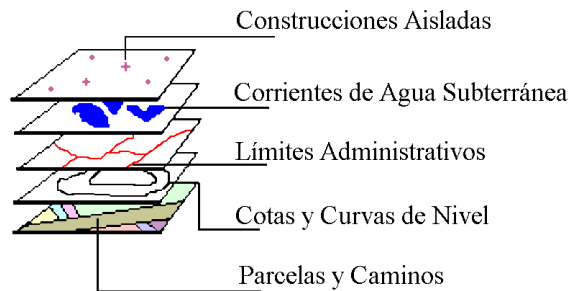
Fig. 2.1 Ejemplo de un mapa creado en 1679

Para localizar información importante se tenían que manipular tablas estadísticas impresas junto con mapas de regiones, con esto se originaba un proceso lento, además de proporcionar resultados erróneos y falta de actualización en los datos.

## 2.2. Origen y desarrollo de GIS

Los Sistemas de Información Geográfica (GIS, en inglés: Geographical Information System), en la actualidad está ampliamente difundido tanto en la geografía como en otras ciencias, en especial aquellas vinculadas con la planificación territorial y la resolución de problemas socioeconómicos y ambientales.

Para cualquier ciencia relacionada con el espacio, en especial la geografía, analiza el territorio a través de distintas capas temáticas (el suelo y sus usos, red hidrográfica, carreteras, caminos, etc.) usando y detectando, aislada o conjuntamente distintas capas de información en la misma zona [Cebrian, 1988<sup>33</sup>], ver figura 2.2.



**Fig. 2.2** Distribución de las diferentes capas de información

De esta forma, el investigador puede analizar cada una de estas capas temáticas dependiendo de los objetivos de estudio. En este sentido, la gran ventaja de los GIS's es que puede relacionar las distintas capas entre sí, lo que le concede a estos sistemas una espectacular capacidad de análisis, pudiendo responder a peticiones complejas y, por ello, producir mapas derivados que pueden representar situaciones reales, escenarios hipotéticos o simulados de gran utilidad [Cebrian, 1988<sup>33</sup>].

El desarrollo de los GIS's comienza con el cambio del formato analógico (cartografía convencional) al digital. La aplicación de las computadoras al análisis de la información espacial en la década de los años 60's. Sin embargo hasta principios de los años 80's se mantuvieron en fase de investigación [CESGA<sup>3</sup>].

El uso de los GIS's ha aumentado enormemente en las décadas de los ochenta y noventa; ha pasado del total desconocimiento a la práctica cotidiana en el mundo de los negocios, en las universidades y en los organismos gubernamentales; a partir de 1990 la disponibilidad de información GIS para entornos Windows y el abaratamiento de los equipos de cómputo ha propiciado su aplicación en empresas

privadas, que lo utilizan como usuarios finales mas que como productores de información.

Hoy en día diversas organizaciones invierten grandes sumas de dinero en el desarrollo de GIS; es previsible además que durante los próximos años se inviertan miles de millones más [CESGA<sup>3</sup>].

Un GIS es una herramienta propia de una organización, no una herramienta personal, es un gran sistema informático cuya implantación en una organización es siempre gradual y costosa. Se requiere siempre la adecuación del sistema al trabajo requerido, mediante programación (frecuentemente realizada por el suministrador del GIS) y recopilación de los datos necesarios (suministrados por otras organizaciones o introducidos por el cliente), junto con los profesionales del área de estudio en cuestión [CESGA<sup>3</sup>].

Es lógico, por tanto, que hayan sido propuestas varias definiciones; una definición precisa y completa podría ser:

“Un sistema GIS es un conjunto de hardware, software, datos geográficos, personas y procedimientos organizados para capturar, almacenar, manejar y desplegar eficientemente rasgos de información referenciadas geográficamente” [Levachkine 2000<sup>1</sup>].



**Fig. 2.3** Elementos que componen a un GIS

Actualmente es difícil encontrar en los medios administrativos, militares, científicos o económicos una tarea en que un GIS no sea una herramienta poderosa para representar y/o analizar información; sin embargo, su creación y aplicación están atrasados respecto a las necesidades prácticas [Levachkine 2000<sup>1</sup>].

En vista de que los GIS's tienen gran importancia, existe la necesidad de producir nuevos mapas digitales y los que ahora están en papel, se integren a los modernos sistemas automáticos.

Los planos que contienen mapas se pueden encontrar en una gran variedad de escalas que pueden ir desde 1:1,000,000 hasta 1:20,000 y aún mayores; el mapa contiene básicamente, objetos puntuales, poligonales, lineales y de texto:

- Los objetos puntuales son símbolos que contienen tamaño y representación similares entre su mismo tipo para reconocerlos, como son ciudades, aeropuertos, yacimientos, etc.
- Los objetos lineales, no conservan un patrón específico, poseen forma típica, color, largo indeterminado y espesor mínimo, pueden representar la red de carreteras, la red ferroviaria e hidrográfica, la de los límites estatales e internacionales, etc.
- Los objetos poligonales dependen de la escala del mapa, tienen área no determinada, en su interior contienen fondo rayado, colado o marcados de manera distinta a los demás, los polígonos representan zonas económicas, con alguna producción agrícola determinada, clima, etc.
- El texto nos proporciona información concerniente a algún objeto específico que se encuentre en el mapa, sea en la leyenda o traslapado con los mismos.

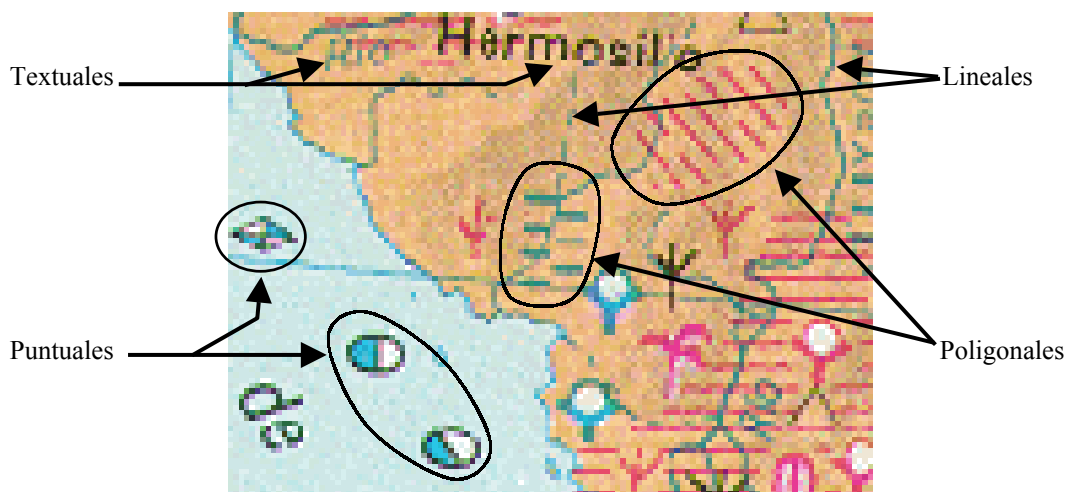


Fig. 2.4 Diferentes formas de representar los objetos en un mapa temático.

Los mapas son el producto de cuadrillas de cartógrafos y topógrafos entre otras disciplinas, que salen al campo para situar y describir gráficamente los hechos naturales y aquellos creados por la mano del hombre.

### 2.3. Representación de datos en un GIS

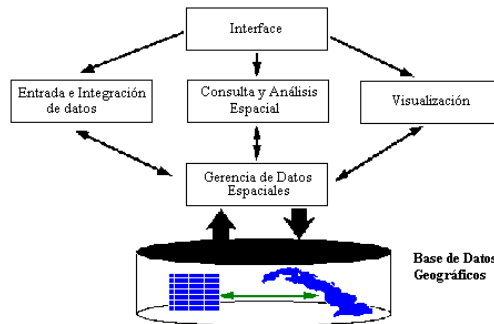
La visión que tiene un GIS del mundo real está dividida en distintas variables temáticas, cuyos valores se pueden estimar en cualquier lugar; cada variable es un estrato o capa de la base de datos. En cada estrato los datos tienen los mismos componentes conceptuales.

Según esta representación, el espacio está constituido por elementos cuya localización precisa es posible por su referenciación a un sistema de coordenadas, además, los elementos espaciales pueden tener atributos temáticos.

En los GIS's se trabaja con dos tipos de datos:

- Datos espaciales, estos a su vez, pueden desglosarse en otras dos componentes:
  - Geométricas; coordenadas (x, y).
  - Topológicas; que hacen referencia a las relaciones espaciales con los diferentes elementos entre sí.
- Atributos temáticos

**En un GIS sus fuentes de información son básicamente mapas en papel, fotografías aéreas, imágenes de satélite y dibujos esquemáticos, en ellos se tiene que transformar la información de imágenes raster a vectoriales.**

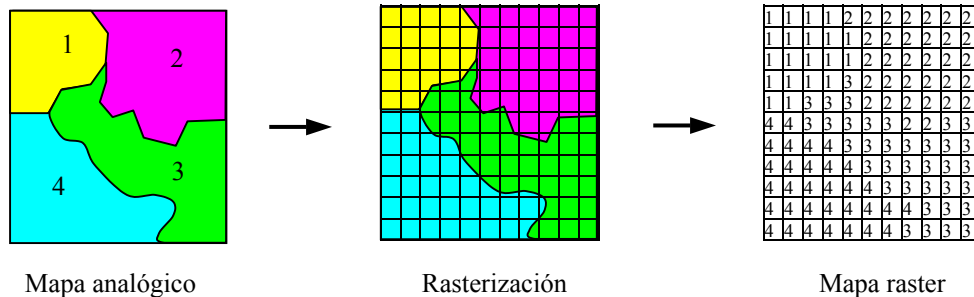


**Fig. 2.5** Arquitectura de un GIS

Según sea la forma de organizar estas componentes se puede hablar de dos tipos de modelos GIS, con ventajas e inconvenientes que determinan el uso de uno o de otro, dependiendo de la naturaleza de la investigación (métodos, escalas, objetivos, etc.), como son, el modelo raster y vectorial.

### 2.3.1. Modelo raster

El modelo raster representa la realidad a través de superficies determinadas que queda dispuesta en forma de matriz, en la que cada elemento está representado por un pixel, véase la figura 2.6.



**Fig. 2.6** Cambio del modelo analógico al digital (raster)

De esta manera, cada celda tiene un valor que la identifica y/o agrupa con las demás, lo que resulta útil para representar fenómenos que se manifiestan en el territorio, como por ejemplo:

- Variables físicas, como son la topografía y precipitaciones.
- Regiones, que estarían constituidas por píxeles agrupados por el mismo valor.
- Emisiones y reflexiones de energía, que son captadas por sensores de satélite.

Las celdas sólo pueden tener un valor, por lo que los objetos representados tienen que estar dispuestos en diferentes archivos, según la variable que se esté tratando. Esta es una de las limitaciones del modelo raster, ya que son necesarias muchas más capas que en un modelo vectorial.

En la codificación de los datos raster se deben tener en cuenta dos factores:

- La localización, que viene determinado por el lugar que ocupan dentro de la matriz, y
- El valor, que viene representado por un número que debe tener un significado establecido.

En el almacenamiento de los archivos raster no es necesario que se incluyan las coordenadas, únicamente los valores. De cualquier manera la cantidad de celdas que son necesarias para obtener una resolución significativa es elevada [Bernhardsen, 1992<sup>35</sup>].

Los modelos raster resultan útiles para el tratamiento de las imágenes de satélites y especialmente en estudios del impacto ambiental, ya que es posible traducir las escenas reales a formato digital con “un antes” y “un después”, sobre todo en aquellos fenómenos que son sólo perceptibles mediante sensores [López, 2000<sup>34</sup>].

### 2.3.2. Modelo vectorial

La representación espacial de un modelo vectorial se basa en la localización de puntos individuales según determinadas coordenadas, por lo que se pueden representar puntos, líneas, polígonos, etc.; los elementos están representados gráficamente y a su vez mantienen un vínculo con una base de datos.

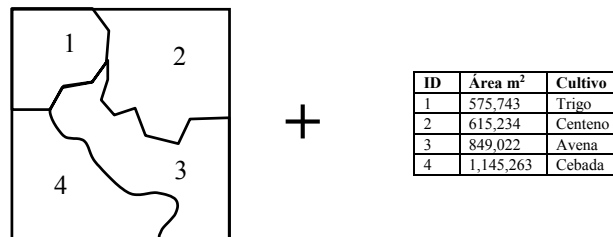


Fig. 2.7 Componentes del modelo vectorial: Gráfico más base de datos.

En la tabla de la base de datos, cada uno de los elementos tendrán asignados sus parámetros identificativos dentro de un registro. Normalmente un GIS mantiene varias tablas relacionadas dentro de una base de datos.

De esta manera se consigue llevar hasta su máxima consecuencia la definición de un GIS, como una herramienta que aplica la operatividad de una base de datos a la representación cartográfica de sus registros, pero que además toda modificación en su componente gráfica tendrá un efecto correspondiente en la base de datos.

**La mayor ventaja estriba en la posibilidad de asignar varios atributos a un mismo elemento dentro de una misma cobertura, e incluso la posibilidad de asociar varios temas dentro de cada una [López, 2000<sup>34</sup>].**

Entre los modelos vectoriales se distinguen los modelos “espagueti” y los topológicos:

- Los primeros se limitan a identificar sus elementos según las coordenadas; las fronteras comunes entre polígonos se duplican, ya que es la única forma de mantenerlos como unidades independientes.
- El modelo topológico tiene en cuenta, además de su localización por coordenadas, las relacionan con los elementos adyacentes, esta relación se mantiene aunque se altere la forma y la extensión de los elementos, lo que significa una ventaja a la hora de manipular los datos.

Estos modelos presentan ventajas sobre los raster al momento de analizar redes y realizar análisis de contigüidad y conectividad. La superposición de coberturas vectoriales tiene también más posibilidades de análisis, ya que mantiene la forma y los elementos de las coberturas, mientras que en los modelos raster se limita a la superposición de matrices que dan nuevas coberturas; sin embargo, como se mencionó anteriormente, los modelos raster se acoplan mejor al tratamiento de las imágenes de satélites [López, 2000<sup>34</sup>].

## 2.4. Procesamiento digital

Durante años las compañías han creando planos cartográficos manualmente, así acumularon grandes volúmenes de información que no pueden tener acceso directamente por una computadora, por lo que se ven obligadas a convertirlas al formato vectorial de forma manual [Levachkine 2000<sup>1</sup>].

Para vectorizar los mapas que están en papel, se utilizan mesas digitalizadoras, este método requiere de una labor muy intensa y lenta que demanda de una gran interacción humana, esto conduce a errores por el cansancio visual; un mapa promedio contiene más de 10,000 elementos lineales [Levachkine 2000<sup>1</sup>]; posteriormente se utilizan en un GIS que permite acceder a la información más eficazmente, de esta manera se abaten los costos de manejo de la misma.

El uso del escáner y los algoritmos de reconocimiento de patrones, establecen una alternativa prometedora para un análisis completo del mapa en forma automática o semiautomática. Estos modelos esencialmente deberán ser capaces de discriminar los elementos representados en el mapa y transformarlos a un formato estandarizado. Aunque esta idea no es nueva, las soluciones que minimicen la interacción humana aún se encuentran bajo investigación [Levachkine 2000<sup>1</sup>].

El problema principal para desarrollar la digitalización automática estriba en la gran cantidad de capas que contiene cada mapa, donde están parcialmente ocultas unas por otras, esto sólo se ha logrado solucionar parcialmente.

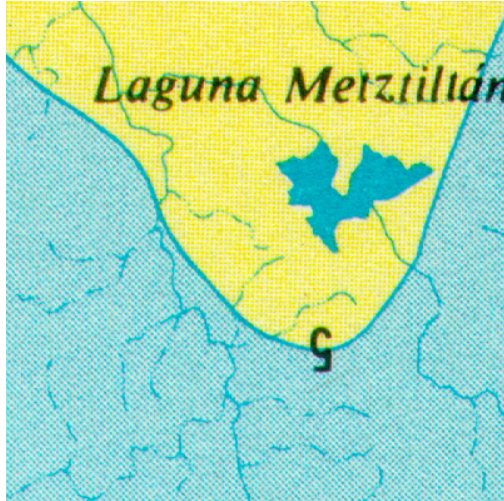
Los planos cartográficos a color están formados por varias capas de información que traslapan sus datos con los demás, en sus colores se pueden confundir la terminación de un elemento que pertenece a una capa con el inicio de otro elemento, que forma parte de otra capa [Levachkine 2000<sup>1</sup>].

De acuerdo a la manera de convertir los mapas en papel a vector se pueden agrupar en cuatro conjuntos: manual, interactivo, semiautomático y automático [Levachkine 2000<sup>4</sup>, Levachkine 2000<sup>5</sup>].

- En los sistemas manuales, el operador utiliza el ratón para marcar la localización de los puntos que conforman las líneas y teclea la información de sus atributos para que sea guardada en una tabla.
- Para los sistemas interactivos, el analista le indica al sistema donde inicia un rasgo lineal, así el programa puede seguir esta línea, pero si es interrumpida o existe una bifurcación, por cualquier causa, el operador debe señalar nuevamente donde debe continuar, y esto se deberá hacer para todas y cada una de las líneas o segmentos que conforman el plano.
- Para los sistemas semiautomáticos aparecieron algunos correctores de vectores, como el R2V (Automated Raster to Vector Conversion Software for Mapping and GIS), que permiten reducir las horas hombre de digitalización de 1.5 a 2 veces.
- Para los sistemas automáticos sólo permite reconocer un número limitado de elementos cartográficos y posteriormente el usuario tiene que corregir los errores que se generaron y completar la tarea inconclusa, que en algunos casos sobrepasa a la tarea automatizada.

La extracción de rasgos lineales es normalmente realizada manualmente por varias razones, seguido aparecen varios tipos de problemas dentro de un mapa raster con las siguientes características:

- El proceso de escaneo es la principal fuente de errores, éste introduce ruido Gaussiano que distorsiona el color de la imagen raster, provocando que el color de las líneas no sea constante en toda su trayectoria.
- Elementos de otras capas que pueden tener el mismo color, estos se tocan, se cruzan y se traslapan.
- Los colores cambian dependiendo del color de la región que atraviesan.



Características de esta imagen:

- El río es de tono verdoso en la región de color amarillo.
- Río es de tono azulado en la región de color azul difuminando.
- El límite de las regiones también es de color azulado.
- Existen cruces del río con el límite de las regiones del mismo color.
- Los ríos están compuestos por líneas parcialmente discontinuas.

**Fig. 2.8** Características de los colores en un mapa raster

Aunque los segmentos de rasgos lineales continuos, sin ramificaciones o cruces, pueden ser detectados y vectorizados por procesos automáticos, los métodos manuales son frecuentemente utilizados para este fin y unir segmentos donde las líneas son discontinuas dentro de una entidad continua, por ejemplo, un tramo de río que va por debajo de la tierra, se dibuja de forma discontinua o no se dibuja.

La conversión de información impresa en papel a datos "inteligentes" para computadora, ha sido el tópico de investigación intensiva en la última década [Levachkine 2000<sup>1</sup>].

La investigación enfocada a la interpretación de planos cartográficos tiene todavía varios campos que tratar, dentro de estos destaca la segmentación, la identificación y la separación de la capa de líneas.

## 2.5. Propósito del proyecto

El propósito de un proyecto para el procesado de planos cartográficos deberá consistir en proporcionar una interpretación totalmente automática, con llamadas al operador sólo cuando el programa detecte casos sospechosos. Las dificultades principales de esta tarea son debidas a la densidad del mapa y a la variedad de leyendas. Se puede adoptar la estrategia de "pelar la cebolla" donde se eliminan del mapa las capas ya identificadas para facilitar la interpretación de las restantes más complejas [Deseilligny 1997<sup>7</sup>].

Para realizar este tipo de proyecto es recomendable desarrollar un sistema de bajo costo con poca intervención manual. No se considera la vectorización interactiva o la interpretación automática con verificación humana exhaustiva como económicamente

viable. Por el otro lado, un sistema completamente automático libre de error no parece ser realista [Schavemaker 1997<sup>6</sup>].

En esta tesis se desarrollarán métodos que trabajen en forma semiautomática y automática, es importante que el proceso sea completo, para minimizar las intervenciones del usuario que en la mayoría de los casos es costoso.

Vamos a extraer los elementos lineales que pertenecen a una capa, previamente seleccionada y separar esa capa del resto (una a la vez), posteriormente se deberá restaurar a las líneas que se extraigan incompletas, se considera importante la información restante del mapa para este proceso, éstas fueron depositadas anteriormente en otra capa en el proceso de extracción.

Las líneas están asociadas a los elementos que conforman el mapa, deberán ser identificadas primeramente por su color similar, se podrán encontrar aisladas, tocando o inclusive traslapadas por otras zonas; una vez identificadas plenamente se procede a su extracción para que pase a formar parte de su propia capa. Las “rupturas” o “vacíos” que tienen las líneas al ser ocultadas por otros objetos se restauran y se dejan completas a las mismas.

- Se van a utilizar técnicas de Procesamiento de Imágenes Digitales para poner el mapa raster en condiciones adecuadas para su procesamiento.
- Técnicas de Reconocimiento de Patrones para la clasificación de las líneas de interés, apoyándose en un análisis de color.
- Para su segmentación o extracción se utilizarán técnicas de Visión por Computadora.
- Posteriormente se procederá a su reconstrucción, basándose en las características locales, tanto de las líneas como de los demás objetos presentes en el mapa.

Lenguaje y equipo utilizado:

- El lenguaje es C++, del paquete comercial de C++ Builder 3®.
- El sistema operativo es Windows 98®.
- Una computadora Pentium II® a 266 MHz.
- 128 Mb en Memoria RAM.

El tiempo de proceso a una imagen de tamaño 2,500 por 3,000 pixeles depende de la complejidad de la misma.

® Marcas Registradas por sus respectivos fabricantes.