A Bilingual Course!

The Principles and Applications of GIS

地理信息系统原理与应用

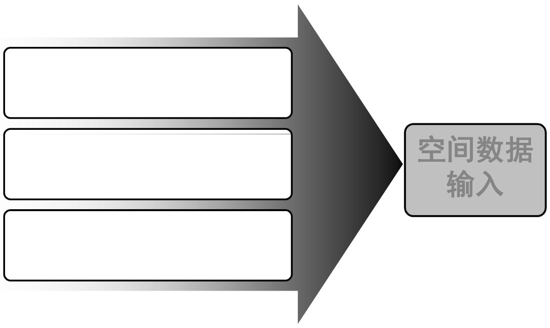
Presented by ***Hu Cheng***
School of Environmental Studies, China University of Geosciences

---

Chapter3 – part 2

Building a Geospatial Database
——Spatial data input & quality control

Primary Tasks of Spatial Data Input



二、Spatial data input

1. Characteristics and requirements
2. Source spatial data
3. Approaches of spatial data input
4. Post-Processing of data input

三、Spatial data quality

1. Where is quality problem from?
2. Representative spatial data error types
3. Spatial data quality control

二、 Spatial data input

1、数据输入特点及要求

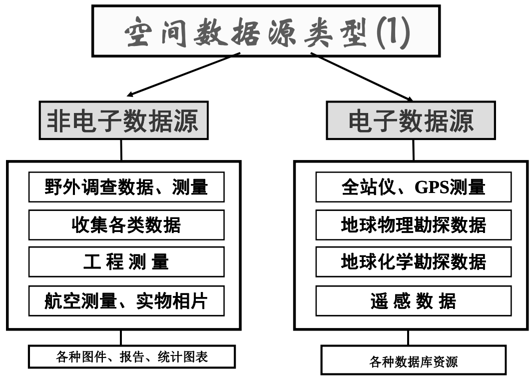
- 数据输入与输出功能是GIS与外部世界通讯（交流）的工具
- 数据输入是将数据编码并写入GIS数据库的过程
- 创建一个精确数字化的数据库是GIS运用最关键的所在。

■ 数据输入由几何数据和属性数据输入两种类型构成

■ 几何数据：记录空间实体的位置、拓扑关系和形状特征；

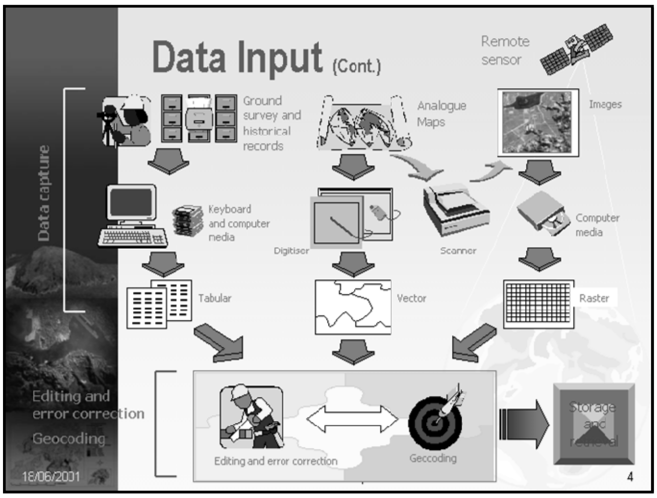
■ 属性数据：表示空间实体具有的各种性质。

2、数据来源



空间数据源类型(2)

- 基础制图数据：自然地理数据（地形地貌、河流、地质背景、社会、经济人口等）
- 自然资源数据：如各种资源的空间分布、储量、品位
- 调查统计数据 不同专项调查数据
- 数字高程数据 DEM
- 行业法规、技术、法律文档数据
- 已有系统数据 其他不同类型的数据库资源
- ……



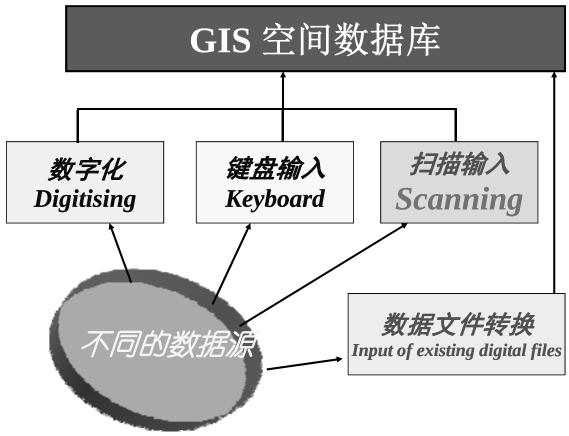
3、Main approaches

1）、键盘输入（Keyboard entry）

2）、手工数字化（Manual digitising）

3）、扫描（Scanning）

4）、已有数据的输入（Input of existing digital files）



3.1 Using Keyboard

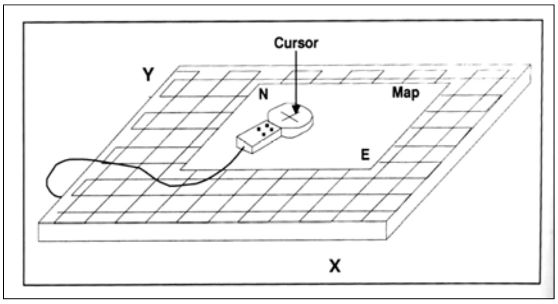
- 主要输入属性数据，很少输入定位数据
- 配合手工数字化输入定位数据（对于数据量较小、并且已知地物精确坐标的情况下，可以采用键盘录入。（输入坐标值））
- ArcGIS中的COGO 扩展模块常用此种方法

3.2 Manual digitising

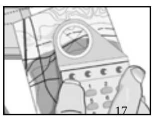
- 数字化是录入定位数据的最常用方法。
  数字化图形输入板（digitizing tablet）
  简便数字化仪（simply digitizer）
- 数字化仪= 一块面板+一个定位设备组成。定位设备称作数字化光标，用来追踪空间目标的位置。

是一项单调而细致的工作！

数字化仪结构示意图



- 数字化面板内部是由金属丝组成的精细网格。
- 垂直金属丝记录着X坐标，而水平金属丝记录Y坐标。
- 数字化仪坐标的范围取决于金属丝网的密度：即数字化仪的分辨率。数字化仪分辨率的范围通常在200 LPI（线/英寸）到2000 LPI 之间。
- 数字化光标上有一个嵌有十字丝的透明窗口。当用户按下某个键时，就会传送一个电磁信号，并且十字丝所在的位置即被某根垂直及水平金属丝记录下来。



- 数字化仪可以用来输入矢量数据，例如通过记录节点的坐标来输入点、线和多边形的边界。
- 几款数字化仪品牌：
  (1). Calcomp
  (2). Summagrid
- 在设定数字化仪时通常要考虑到以下6种参数：
  A.Baud rate(波特率) 150,300,600,1200, 2400, 4800, 9600, 19200 ...
  B.Parity (奇偶性) odd, even
  C.Data Bits (数据位)7, 8
  D.Stop Bits (停止位)
  E.Point Mode (点模式)
  F.Stream Mode (流线模式)

Summary:

Data Media Conversion--*Digitizing: manually* tracing a map or aerial

- Applied to map or aerial photo
- Use hard copy map/photo on table/tablet, **or** scanned image on screen (heads-up digitizing)
- pen or cursor detects x, y coords
- coordinates are in inches/cms from lower left (0,0)
- control points (tic marks) relate digitized coordinates to real world lat/long coordinates
- coordinates captured in *stream* or *point* mode
- accuracy of table (but not user!) usually better than 0.1 mm
- all nodes and polygons should be marked and numbered first
- essentially a vector approach

Problems:

- paper maps unstable
  - crease and fold
  - stretch with humidity (up to 3%)
  - photos more stable (0.2%)
- map errors transferred to GIS
  - maps often prepared for display not accuracy
- human hand very shaky
- often generates undershoots, overshoots, & double lines
  - editing and clean-up essential

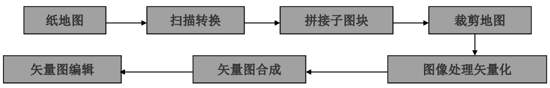
3.3 Scanning

- 利用扫描仪记录原始（模拟）图像对应部分的反射值，形成栅格（数字）矩阵或数字图像(image)。
- 遥感数据也是利用某种传感器扫描地球的表面而获得的。（比如：SPOT 5卫星HRG (High Resolution Geometry)传感器）
- 扫描仪的色彩分辨率取决于每个像素的数值范围，8位=256色，…
- 扫描仪的空间分辨率以每英寸记录点的数量表示（dpi=dot per inch），从100 dpi到1200 dpi
- 扫描输入是GIS建立空间数据库最主要的手段之一。

- **扫描输入后的图像一般需要进行矢量化。**

通常是将地图扫描后，作为底图显示在屏幕上，用鼠标参照底图进行采点。由于鼠标定位不如数字化仪精确，所以一般用于输入一些简单图形。

扫描矢量化的主要步骤：



Usually vectorize using conversion software !

Summary:

Data Media Conversion--*Scanning: automated* recording of map or aerial

- **Produces "dumb" raster data**
  - vectorize using conversion software
  - Create "smart" image using digital image processing techniques
- **electromechanical**
  - $100-$50,000 instruments
  - drum or flatbed
  - scan resolution depends on price!
    - » down to 20 microns (millionth of m)
- **Scanners v. sensors**
  - Sensors collect data directly in digital form (e.g. digital cameras)
  - Sensor resolution now (2005>) matches that of photos, so scanning photos becoming old technology
  - Still lots of paper maps around e.g. property ownership records
- Great if need only raster representation
- Automated creation of vector data from scanning very problematic:
  - docs must be clean
  - complex line work adds error
  - lines shouldn't be broken with text.
  - text may be interpreted as lines
  - automatic feature detection (road versus railroad) difficult
- ESRI's *ArcScan for ArcGIS* (included with ArcEditor) provides interactive, semi-automated raster to vector conversion.
  - Other vendors offer specialized conversion software
- Digital image processing techniques used to create "smart raster"
  - Identify feature type within each raster

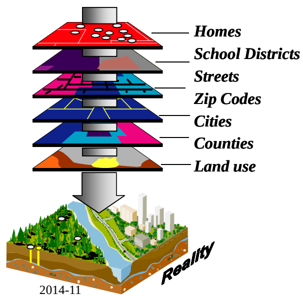
Calcomp - 436CX PLUS

- 平板扫描仪
- 滚筒扫描仪

Noticeable issues during data input (1)

- 分层采集：

在进行数字化之前，首先要确定需要数字化哪些信息，GIS系统对空间数据往往采用分层管理，所以要确定输入哪些图层，以及每个图层包含的具体内容。



Noticeable issues during data input (2)

- 坐标系统、空间位置确定与匹配：
  确定坐标系统、投影类型
  两次输入之间地图的位置匹配
  不同时间输入，图件与数字化板可能相对于发生错动，导致前后两次录入的坐标就会偏移或旋转。输入至少三个定位点(Tick Marks),进行定位点坐标之间的关系进行匹配。
- 选择输入模式：
  点模式(point mode)
  流模式(stream mode)
  距离流模式(Distance stream)
  时间流模式(Time stream Mode)

Noticeable issues during data input (3)

- 距离流模式(Distance stream)

当前接收的点与上一点距离超过一定阈值才记录该点坐标值

- 时间流模式(Time stream Mode)

按照一定时间间隔对接收的点进行采样

  – 时间流模式录入的优点：当录入曲线比较平滑时，录入人员往往移动游标比较快，这样记录点的数目少；而曲线比较弯曲时，游标移动较慢，记录点的数目就多。
  – 采用距离流模式，容易遗漏曲线拐点，从而使曲线形状失真。所以在保证曲线的形状方面，时间流方式要优于距离流方式。

3.4 Input of existing digital files

- 购买和使用已有的数据资料是GIS数据输入的最有效途径。
- 常用数据可以从有关的政府和商业部门获得。
- 栅格数据导入的最常用文件格式有：

  TIF; JPG; JPE; BMP; GIF; BSQ; BIP;BIL,TM(Landsat); SPOT *etc.*
- 矢量数据导入的最常用文件格式有

  DXF；DWG; E00；SHP, *etc.*
- 数字文件转换输入包括：
  (1) 数据转移和存储介质转化
  (2) 数据格式的转化
  (3) 动态数据交换和数据库共享
  (4)数据交换标准

Summary:

- In most GIS software data is organized in themes as data layers.
- Data to be input as separate themes and overlaid based on analysis requirements.
- This can conceptualized as vertical layering the characteristics of the earth's surface.
- The overlay concept is so natural to cartographers and natural resource specialists that it has been built into the design of most CAD vector systems as well.
- This concept is also used to logically order data in most GIS software.
- The terminology may differ between GIS software, but the approach is the same.
- A variety of terms are used to define data layers in commercial GIS software.

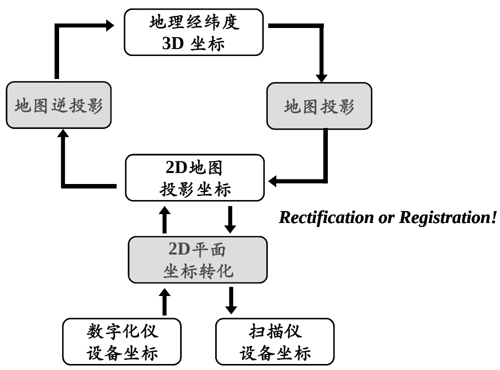
  themes coverages layers levels objects ……
- Data layer and theme are the most common and the least proprietary to any particular GIS software and accordingly, as used throughout the book.
- In any GIS project a variety of data layers will be required. These must be identified before the project is started and a priority given to the input or digitizing of the spatial data layers.
- This is mandatory, as often one data layer contains features that are coincident with another
- Data layers are commonly defined based on the needs of the user and the availability of data. They are completely user definable.

4、Post-Processing of data input

1. 进行坐标变换或投影变换
2. 图形拼接
3. 修改错误、建立拓扑关系
4. 属性库完善：编码、联结

4.1 坐标变换或投影变换

Transformation and Projection



Georeferencing: Rectification and Registration

providing true earth location/overlaying layers

- *rectification:* rearrangment of location of objects to correspond to a **specific** reference system (usually geodetic)
- *registration:* rearrangment of location of objects of one set so they correspond with those of another, without reference to a specific reference system

Despite formal difference, often used interchangeably

Two methods

- homogeneous transformation via *rotation, translation, scaling, skewing*
  - used for map projection and similar conversions
- differential transformation via *rubber sheeting*
  - used to correctly position distorted images or scanned maps or documents

•Most commonly used to relate images (e.g. scanned photo) to a vector layer, but can also be used to "fix" incorrect positioning of features in a vector layer

•Implemented in ArcMap: via the Georeferencing toolbar for <u>images</u>

via the Spatial Adjustment toolbar for <u>vector layers</u>

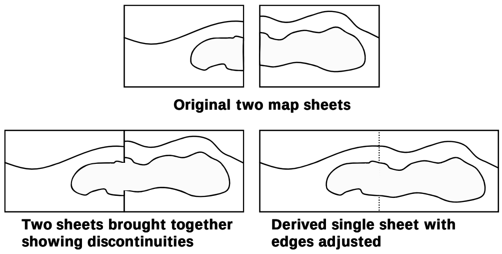
4.2 图 形 拼 接

Edge Matching:

Joining map sheets to create a seamless GIS

- 图形拼接或边缘匹配处理

1) 图幅比较大或者使用小型数字化仪时
2) 输入、输出标准分幅的地形图



4.3 修改错误、建立拓扑关系

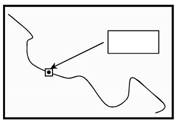
图形数字化常见错误（1）

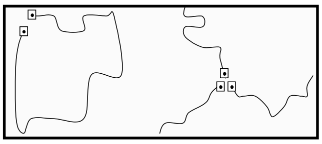
主要原因：空间数据定位的不准确

1）伪节点（Pseudo Node）

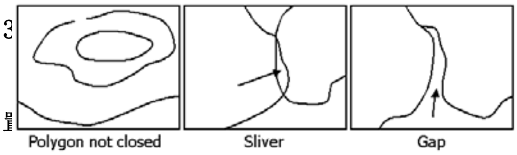
伪节点使一条完整的线变成两段

2）悬挂节点（Dangling Node）





图形数字化常见错误（2）



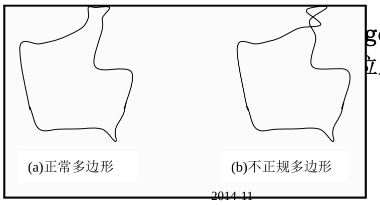
3 ver

线的位

4） gon）

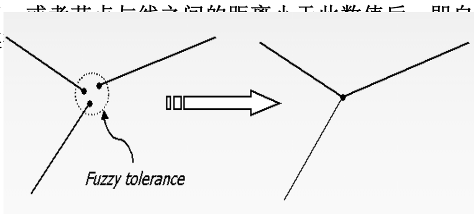
位置不准确引起

的



图形数字化错误修正

- 上述的错误，一般在建立拓扑的过程中都会发现，而且需要手工编辑修改;
- 悬挂节点的问题则可以由软件系统自动修改。通常的实现办法是设置一个"容许误差 **Tolerance**"，当节点之间 动 连



建立拓扑关系

- 在图形错误修改完毕之后，可以建立图形数据的拓扑关系。拓扑关系的建立通常由计算机自动生成，目前大多数GIS软件具有完善的拓扑功能；
- 每一次修改数据错误、或增加、减少图形数据后（即对矢量数据进行编辑后），对具有拓扑关系的矢量数据需要重新建立他们之间的拓扑关系

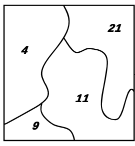
如在ArcInfo 系统中，建立拓扑关系的命令为：
Build point; Build line; build polygon;
Clean point; Clean line; Clean polygon

Summary:

Errors: *detection and removal*

- GIS packages commonly use topological structure checking to detect errors
- Editing based on *node snapping* used to correct errors: moving a feature so its coordinates correspond **exactly** with another's
- snapping conducted based on *tolerances* -- snap if within 1 foot, for example
- ***Care must always be taken to assure that topological "cleaning" does not itself introduce errors (e.g. snapping nodes and lines together which shouldn't be)***

4.4 属性库完善：编码、联结



ID	面积	周长	土壤类型
4	435	880	18
9	210	580	25
11	628	1140	21
21	252	650	15

土壤类型	名称	pH	… …
15	Black soil	6.5	
21	Brown soil	6.0	
25	Red soil	5.0	

在关系数据库中存储空间对象的属性特征值

三 空 间 数 据 质 量

Spatial Data Quality

GiGo: garbage in, garbage out
'Cos it's in the computer, don't mean it's right

It's not the things you don't know that matter, it's the things you know that aren't so.

Will Rogers, Famous Okie GI specialist

"But there are also unknown unknowns: the ones we don't know we don't know." *Donald Rumsfeld*

"Fast is fine, but accuracy is everything." *Wyatt Earp*

空间数据质量

- 空间位置、属性特征和时间是表示实体空间变化的三个基本要属，空间数据是这三者信息的符号记录。而数据质量则是表达它们一致性、完整性和统一性的程度。
- 空间数据是对现实世界的抽象表达，由于现实世界的复杂性和模糊性，由于人类认识水平和表达能力的限制，这种抽象表达总是有限的。因此，数据质量的好坏是一个相对概念。
- 衡量空间数据质量标准的几个要素（**FGDC**）

（1）位置准确度 — 空间实体的真实位置与坐标数据的接近程度
（2）属性准确度 — 空间实体的属性与其真值相符程度
（3）时间准确度 — 数据更新的时间和频度来衡量
（4）逻辑一致性— 数据结构、数据内容、拓扑关系的一致
（5）数据完整性 — 空间数据覆盖的范围、类型、关系

"精度"与"准确度"区别

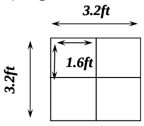
- 精度（**Precision or Resolution**）
  – the exactness of measurement or description
  – Determined by input; can output at lower (but not higher) resolution
- 准确度（**Accuracy**）
  – the degree of correspondence between data and the real world
  – Fundamentally controlled by the quality of the input

Precision or Resolution

*it's **not** the same as scale or accuracy!*

Precision: the exactness of measurement or description

- the "size" of the "smallest" feature which can be displayed, recognized, or described
- Can apply to *space*, *time* (e.g. daily versus annual), or *attribute* (douglas fir v. conifer)
- for raster data, it is the size of the pixel (resolution)
  - e.g. for NT*GIS*C digital orthos is 1.6ft (half meter)
- raster data can be resampled by combining adjacent cells; this decreases resolution but saves storage
  - eg 1.6 ft to 3.2 ft (1/4 storage); to 6.4 ft (1/16 storage)

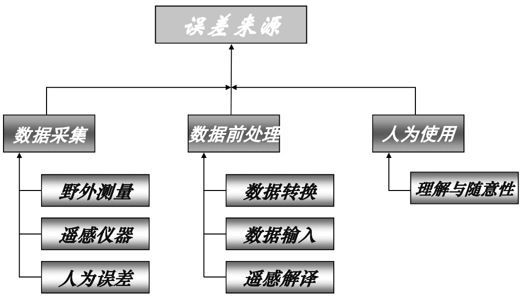


- resolution and scale
  - generally, *increasing* to larger scale allows features to be observed better and requires higher resolution
  - but, because of the human eye's ability to recognize patterns, features in a *lower* resolution data set can sometimes be observed better by *decreasing* the scale (6.4 ft resolution shown at 1:400 rather than 1:200)
- resolution and positional accuracy
  - you can see a feature (resolution), but it may not be in the right place (accuracy)
  - higher accuracy generally costs *much* more to obtain than higher resolution
  - accuracy cannot be greater (but may be much less) than resolution (e.g. if pixel size is one meter, then best accuracy possible is one meter)

Scale, Resolution & Accuracy in GIS Systems

- On paper maps, ***scale*** is hard to change, thus it generally determines ***resolution*** and ***accuracy***--and consistent decisions are made for these.
- A GIS is *scale independent* since output can be produced at any scale, irrespective of the characteristics of the input data— at least in theory
- in practice, an implicit *range of scales* or *maximum scale* for anticipated output should be chosen and used to determine:
  - **what** features to show
    - » manholes only on large scale maps
  - **how** features will be represented
    - » manhole a polygon at 1:50; cities a point at 1:1,000,000
  - appropriate levels for **accuracy and precision**
    - » Larger scale generally requires greater resolution
    - » Larger scale necessitates a higher level of accuracy
- GIS also helps with the *generalization* problem implicit in paper maps
  - A road drawn with 0.5 mm wide line (the smallest for decent visibility)
    - » At 1:24,000 implies the road is 12 meters (36 feet) wide
    - » At 1:250,000 implies the road is 125 meters (375 feet) wide
  - At least in a GIS you can store the true road width, but be careful with plots!

3.1空间数据质量问题的来源



空间数据质量问题的来源

数据处理过程	误 差 来 源
数据搜集	野外测量误差：仪器误差、记录误差 遥感数据误差：辐射和几何纠正误差、信息提取误差 地图数据误差：原始数据误差、坐标转换、制图综合及印刷
数据输入	数字化误差：仪器误差、操作误差 不同系统格式转换误差：栅格-矢量转换、三角网-等值线转换
数据存储	数值精度不够 空间精度不够：每个格网点太大、地图最小制图单元太大
数据处理	分类间隔不合理 多层数据叠合引起的误差传播：插值误差、多源数据综合分析误差， 比例尺太小引起的误差
数据输出	输出设备不精确引起的误差 输出的媒介不稳定造成的误差
数据使用	对数据所包含的信息的误解 对数据信息使用不当

Sources of Error

Error is the inverse of accuracy. It is a discrepancy between the coded and actual values.

Sources

- Inherent instability of the phenomena itself
  - E.g. Random variation of most phenomena (e.g. leaf size)
- Measurement
  - E.g. surveyor or instrument error
- Model used to represent data
  - E.g. choice of spheroid, or classification systems
- Data encoding and entry
  - E.g. keying or digitizing errors
- Data processing
  - E.g. single versus double precision; algorithms used
- Propagation or cascading from one data set to another
  - E.g. using inaccurate layer as source for another layer

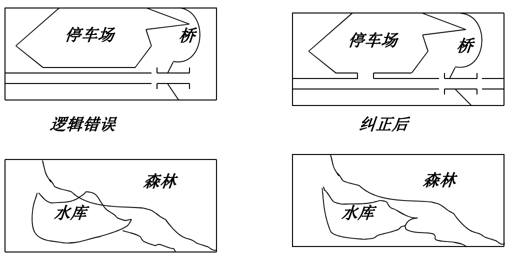
Example for Positional Accuracy

- choice of spheroid and datum
- choice of map projection and its parameters
- accuracy of measured locations (surveying) of features on earth
- media stability (stretching ,folding, wrinkling of maps, photos)
- human drafting, digitizing or interpretation error
- resolution &/or accuracy of drafting/digitizing equipment
  - Thinnest visible line: 0.1-0.2 millimeters
  - At scale of 1:20,000 = 6.5 - 12.8 feet (20,000 x 0.2 = 4,000mm = 4m = 12.8 feet)
- registration accuracy of tics
- machine precision: coordinate rounding error in storage and manipulation
- other unknown

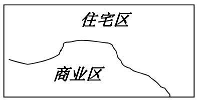
空间数据的误差类型

- ♉ 几何误差（Positional Accuracy, Quantitative accuracy)
- ♉ 属性误差（Attribute Accuracy，Consistency)
- ♉ 时间误差（ Temporal Accuracy)
- ♉ 逻辑误差（ Logical Consistency)

1）逻辑误差

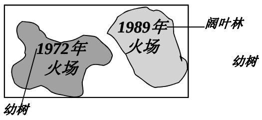


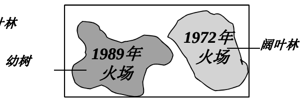
2）属性错误



属性错误　　纠正后

3）时间错误





时间错误　　纠正后

4）几何错误

☺ 点误差

☺ 线误差

Measurement of Positional Accuracy

- usually measured by **root mean square error**: *the square root of the average squared errors*

$$RMSE = \sqrt{\frac{e_1^2 + e_2^2 + e_3^2 + ... + e_n^2}{n-1}}$$

where e_i is the distance (horizontally or vertically)between the tue location of point $_i$ on the ground, and its location represented in the GIS.

- Usually expressed as a probability that no more than P% of points will be further than S distance from their true location.
- Loosely we say that the **rmse** tells us *how far recorded points in the GIS are from their true location on the ground, on average.*
- More correctly, based on the normal distribution of errors, 68% of points will be rmse distance or less from their true location, 95% will be no more than twice this distance, providing the errors are random and not systematic (i.e. the mean of the errors is zero)
  - e.g. for *NTGISC digital orthos* RMSE is 3.2 feet (one meter)
    for *USGS Digital Ortho Quads* RMSE spec. is approx. 33 feet or 10 meters (but in reality much better)

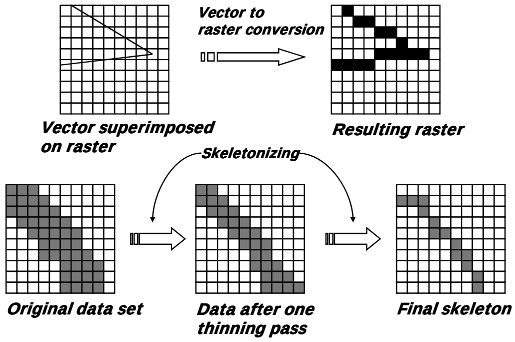
-- with GPS, height is 2 or 3 times *less* accurate in practice at high precision than horizontal (officially the spec is 1.5, but data collection errors affect vertical the most)

3.2 典型的空间数据误差

3.2.1 数据前处理中的误差

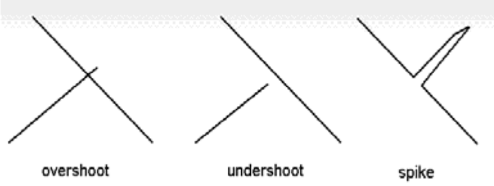
- 数据格式转换（Format conversion）
- 数据输入（ Data Input）
- 边缘匹配（ Edge matching）
- 数据缩编与概化（ Data reduction and generalization ）
- 插值（ Interpolation ）
- 航片解译（ Photo interpretation ）

3.2.2矢量栅格数据转化中的误差



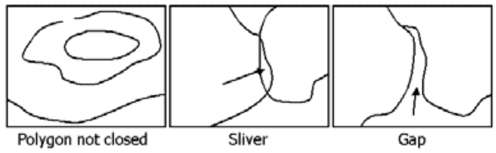
3.2.3 数据输入中的误差

- 地图固定不牢。如果地图发生了移动需要重新输入控制点
- 在数字化的过程中，图纸可能发生伸展或收缩等变形，导致数字化点位偏移了原来位置。
- 用户操作不当导致线段交叉点处的“不足”、“过头”或“细缝”



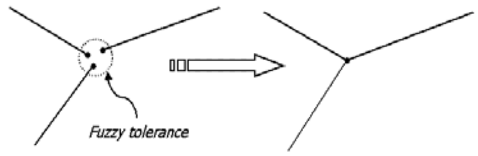
- 空间数据输入过程中常见的误差

polygon not closed; sliver; gap ; attribute errors, etc.



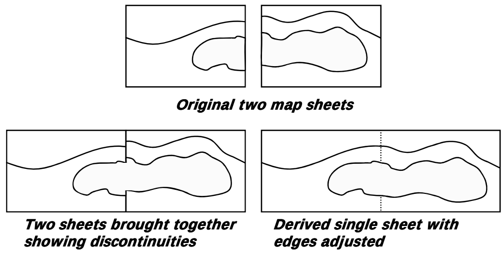
3.2.4 修正过程中的误差

- 设定坐标距离容限值（ fuzzy tolerance）
- 节点连接
- 拓扑关系重建（Re-building topology）

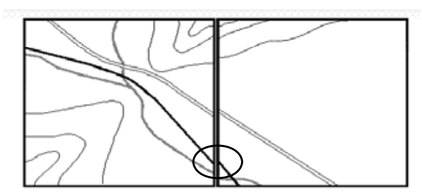


3.2.5 边缘配准中的误差

在自动或手工处理地图的合并过程中，常出现误差



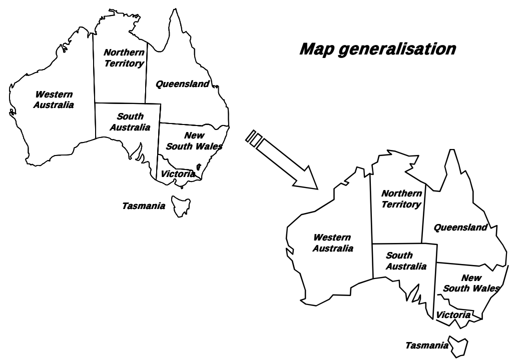
- **跨越地图边界处的目标在综合GIS数据库中常常会导致误差**



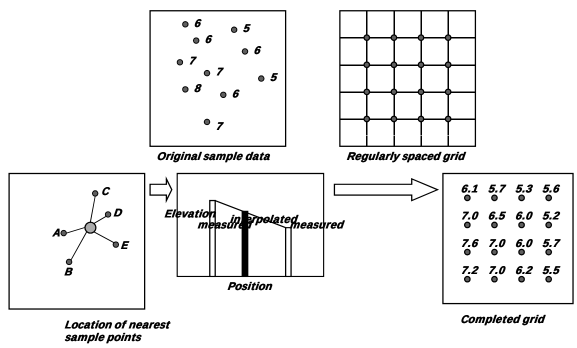
3.2.6 数据压缩和概化

- **数据压缩**（Data reduction）
  (1) 简化数据记录
  (2) 改变比例尺
- **数据概化**（Data generalization ）
  (1) 降低精确度
  (2) 矢 量: 减少节点和多边形
  (3) 栅 格: 用较大的像素尺寸重取样

Map Generalization



3.2.7 插值运算(Interpolation)



3.3 空间数据质量控制

- 应从数据质量产生和扩散的所有过程和环节入手，分别用一定的方法减少误差，常见的方法有：

1. **传统的手工方法：** 质量控制的人工方法主要是将数字化数据与数据源进行比较，图形部分的检查包括目视方法、绘制到透明图上与原图叠加比较，属性部分的检查采用与原属性逐个对比或其他比较方法；
2. 地理相关法：用空间数据的地理特征要素自身的相关性来分析数据的质量。如叠加河流和等高线两层数据时，如河流的位置不在等高线的外凸连线上，则说明两层数据中必有一层数据有质量问题
3. 元数据方法：数据集的元数据中包含了大量的有关数据质量的信息，通过它可以检查数据质量，同时元数据也记录了数据处理过程中质量的变化，通过跟踪元数据可以了解数据质量的状况和变化。

空间数据的元数据

- **在地理空间数据中，元数据(Metadata)是说明数据内容、质量、状况和其他有关特征的背景信息，是描述数据的数据。**

例如：传统的图书馆卡片、出版图书的版权说明、磁盘的标签等都是元数据。纸质地图的元数据主要表现为地图类型、地图图例，包括图名、空间参照系和图廓坐标、地图内容说明、比例尺和精度、编制出版单位和日期或更新日期、销售信息等。用户通过它可以非常容易地确定该书或地图是否能够满足其应用。

元数据的主要作用可以归纳为如下几个方面：

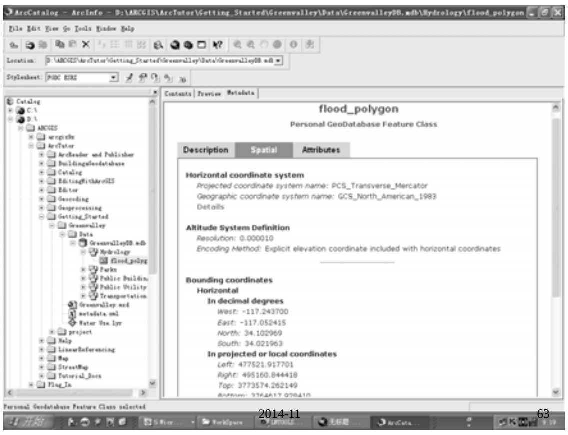
1）帮助数据生产单位有效地管理和维护空间数据、建立数据文档，并保证即使其主要工作人员离退时，也不会失去对数据情况的了解；

2）提供有关数据生产单位数据存储、数据分类、数据内容、数据质量、数据交换网络及数据销售等方面的信息，便于用户查询检索地理空间数据；

3）帮助用户了解数据，以便就数据是否能满足其需求做出正确的判断；

4）提供有关信息，以便用户处理和转换有用的数据。

元数据说明范例



Summary:

Data Quality: *How good is your data?*

- Scale
  - ratio of distance on a map to the equivalent distance on the earth's surface
  - Primarily an output issue; at what scale do I wish to display?
- Precision or Resolution
  - the exactness of measurement or description
  - Determined by input; can output at lower (but not higher) resolution
- Accuracy
  - the degree of correspondence between data and the real world
  - Fundamentally controlled by the quality of the input
- Lineage
  - The original *sources* for the data and the *processing steps* it has undergone
- Currency
  - the degree to which data represents the world at the present moment in time
- Documentation or Metadata
  - data about data: recording all of the above
- Standards
  - Common or “agreed-to” ways of doing things
  - Data built to standards is more valuable since it’s more easily shareable

Summary:

Resolution, Scale, Accuracy & Storage:
illustrating the relationship

Largest (maximum) scale for given pixel size.
Storage is for USGS 7.5 quad. area

(in Texas, USGS quad is about 7 mi x 8.5 mi=60 sq. miles--16 quads for Dallas County)

Source: GPS Technology Corporation

Appendix

FGDC Standards
(status as of March 2004)
For latest, go to:
http://www.fgdc.gov/standards/standards.html

Federal Geographic Data Committee

FGDC: Metadata Standards

Metadata:

- *Content Standard for Digital Geospatial Metadata (version 2.0)* FGDC-STD-001-1998
- *Content Standard for Digital Geospatial Metadata, Part 1: Biological Data Profile* FGDC-STD-001.1-1999
- *Metadata Profile for Shoreline Data* (FGDC-STD-001.2-2001)
- *Content Standard for Digital Geospatial Metadata: extension for remote sensing data* (FGDC-STD-0012-2002)
- *Encoding Standard for Geospatial Metadata* (Draft)
- *Metadata Profile for Cultural and Demographic Data* (dropped)

Current thrust is to integrate FGDC Metadata standards (and other FGDC standards eventually) into International Standards Organization (ISO) standards.

FGDC: Data Accuracy Standard

Geospatial Positioning Accuracy Standard (FGDC-STD-007)

Part 1, Reporting Methodology FGDC-STD-007.1-1998

Part 2, Geodetic Control Networks FGDC-STD-007.2-1998

Part 3, National Standard for Spatial Data Accuracy FGDC-STD-007.3-1998

Part 4: Architecture, Engineering Construction, and Facilities Management (FGDC-STD-007.4-2002),

Part 5: Standard for Hydrographic Surveys and Nautical Charts (Review)

•An umbrella incorporating several accuracy standards.
•Part 3 is the general standard.
•It essentially updates the National Map Accuracy Standard of 1941/47

FGDC: Data Content Standards

- *Cadastral Data Content Standard* FGDC-STD-003
- *Classification of Wetlands and Deep Water Habitats* FGDC-STD-004
- *Vegetation Classification Standard* FGDC-STD-005
- *Soils Geographic Data Standard*, FGDC-STD-006
- *Content Standard for Digital Orthoimagery*, (FGDC-STD-008-1999)
- *Content Standard for Remote Sensing Swath Data*, (FGDC-STD-009-1999)
- *Utilities Data Content Standard*, (FGDC-STD-010-2000)
- *NSDI Framework Transportation Identification Standard*, (Review)
- *Hydrographic Data Content Standard for Coastal and Inland Waterways*, (Review)
- *Content Standard for Framework Land Elevation Data*, (Review)
- *Facility ID Data Standard, (Review)*
- *Address Content Standard (Review)*
- *US National Grid (FGDC-STD-0011-2001)*
- *Earth Cover Classification System, (draft)*
- *Geologic Data Model, (Draft)*
- *Governmental Unit Boundary Data Content Standard, (Draft)*
- *Biological Nomenclature and Taxonomy Data Standard (draft)*
- *National Hydrography Framework Geospatial Data Content Standard (proposal)*
- *Environmental Hazards Geospatial Data Content Standard, (dropped)*
- *NSDI Framework Data layers (under Review—see next slide)*

FGDC: Framework Data Standards

- establish <u>data content</u> requirements for the seven layers of geospatial data that comprise the *National Spatial Data Infrastructure* (NSDI), the base layers needed for any geographic area
  - *geodetic control,*
  - *elevation,*
  - *Orthoimagery*
  - *Hydrography (water)*
  - *Transportation*
  - *Cadastral (landownership)*
  - *governmental unit boundaries*
- *Goals are to*
  - *Facilitate and promote exchange of framework layers between producers, consumers, and vendors thru a common content and way of describing that content*
  - *Lower the cost of data for everyone*
- *For each layer, specifies an integrated application schema in Unified Modeling Language (UML) including feature types, attribute types, attribute domain, feature relationships, spatial representation, data organization, and metadata*
- *no standard specified for data format, but an appendix describes a possible implementation using the Geography Markup Language (GML) Version 3.0, developed through the Open GIS Consortium, Inc. (OGC).*

FGDC: Data Transfer Standards

Spatial Data Transfer Standard (SDTS) FGDC-STD-002

SDTS, Part 1 Logical Specification (FIPSPUB 173-1, July 1994)

SDTS, Part 2 Spatial Features (FIPSPUB 173-1, July 1994)

SDTS, Part 3 ISO 8211 Encoding (FIPSPUB 173-1, July 1994)

SDTS, Part 4 Topological Vector Encoding (FIPSPUB 173-1, July 1994)

SDTS, Part 5 Raster Profile and Extensions (FGDC-STD-002.5, 2000)

SDTS, Part 6: Point Profile, FGDC-STD-002.6, 2000

SDTS Part 7: Computer-Aided Design and Drafting (CADD) Profile (FGDC-STD-002.7, 2000)

•One of the first of the FGDC standards (along with metadata).
•Intended to facilitate transfers between different GIS systems.
•Competitive pressures plus internal weaknesses hindered adoption.

FGDC: Data Symbology and Presentation Standards

- *Digital Geologic Map Symbolization*, (Review)

Unknown

Donald Rumsfeld

- The Unknown As we know,
  There are known knowns.
  There are things we know we know.
  We also know There are known unknowns.
  That is to say We know there are some things We do not know.
  But there are also unknown unknowns,
  The ones we don't know We don’t know.

- 《不知道》

  据我们所知，
  我们已经知道一些，
  我们知道我们已经知道一些，
  我们还知道，
  我们有些并不知道，
  也就是说，
  我们知道有些事情我们还不知道，
  但是，还有一些，
  我们并不知道我们不知道，
  这些我们不知道的，
  我们不知道