

# 空间数据的采集与处理

# （优选）空间数据的采集与处理

# §1 空间数据的来源

- 地图
- 遥感影像
- 统计数据：
  - 国民经济的各种统计数据，如人口数量、人口构成、国民生产总值等等。
- 实测数据：
  - GPS测量数据、地形测量、地籍测量数据等。
- 数字数据
  - 其它GIS中的空间数据或数字地图
- 各种文字报告和立法文件：
  - 如在城市规划管理信息系统中，各种城市管理法规及规划报告

# 地图

- 各种类型的地图是**GIS**最主要的数据源
- 传统地图的缺陷：
  - (1) 地图存储介质的缺陷：存在不同程度的变形，具体应用时，须对其进行纠正。
  - (2) 地图现势性较差: 传统地图更新周期较长，造成现存地图的现势性不能完全满足实际的需要。
  - (3) 地图投影的转换: 由于地图投影的存在，使得对不同地图投影的地图数据进行交流前，须先进行地图投影的转换。

# 遥感影像

- 遥感影像是GIS中一个极其重要的信息源。
- 通过遥感影像可以快速、准确地获得大面积的、综合的各种专题信息，航天遥感影像还可以取得周期性的资料，这些都为GIS提供了丰富的信息。
- 但是因为每种遥感影像都有其自身的成像规律、变形规律，所以对其的应用要注意影像的纠正、影像的分辨率、影像的解译特征等方面的问题。

# §2 空间数据采集

- 空间数据主要包括几何图形数据和属性数据，几何图形数据的获取方法主要有：
  - 大地测量方法
    - GPS测量
    - 全站仪测量
    - 光学经纬仪测量
    - 平板仪测量
  - 摄影测量方法
    - 解析摄影测量、数字摄影测量
  - 遥感方法
  - 现有地形图数字化
  - 数据格式转换

- 对于栅格数据的获取，**GIS**主要涉及使用扫描仪等设备对图件进行扫描数字化，功能比较简单，因为通过扫描获取的标准格式的图像文件，大多可直接进入**GIS**的地理数据库。从遥感影像上直接提取专题信息，需要使用几何纠正、光谱纠正、影像增强、图像变换、结构信息提取、影像分类等技术主要属于遥感图像处理的内容。
- 因此，以下主要介绍**GIS**矢量数据的采集

## 2.1 现有地形图数字化

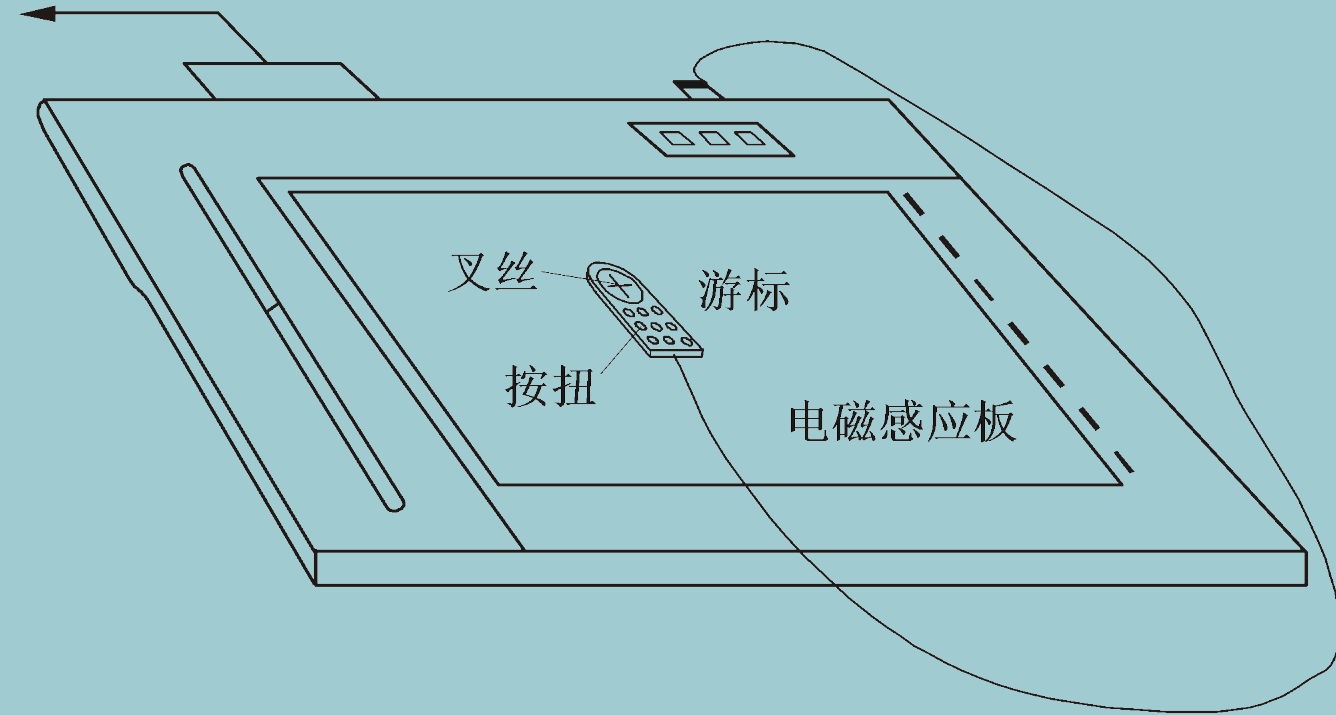
- 现有地形图数字化一般有两种方法：
  - 手扶跟踪数字化
  - 扫描矢量化。



# 手扶跟踪数字化

## 数字化仪

通向计算机接口



- 数字化仪由电磁感应板、游标和相应的电子电路组成。

# 数字化仪工作原理

- 利用电磁感应原理，在电磁感应板的x，y方向上有许多平行的印刷线，每隔 $200\mu\text{m}$ 一条。游标中装有一个线圈。当使用者在电磁感应板上移动游标到图件的指定位置，并将十字叉丝的交点对准数字化的点位，按动相应的按钮时，线圈中就会产生交流信号，十字丝的中心也便产生了一个电磁场，当游标在电磁感应板上运动时，板下的印制线上就会产生感应电流。印制板周围的多路开关等线路可以检测出最大信号的位置，即十字叉线中心所在的位置，从而得到该点的坐标值。

# 数字化过程

- 把待数字化的图件固定在图形输入板上，首先用鼠标器输入图幅范围和至少四个控制点的坐标，随后即可输入图幅内各点、曲线的坐标。
- 通过数字化仪采集数据，数据量小，数据处理的软件也比较完备，但由于数字化的速度比较慢，工作量大，自动化程度低，数字化的精度与作业员的操作有很大关系，所以，目前很多单位在大批量数字化时，已不再采用它。

# 扫描矢量化

- 扫描
  - 扫描参数设置，包括：
    - a、扫描模式的设置（分二值、灰度、彩色），对地形图的扫描一般采用二值扫描，或灰度扫描。对彩色航片或卫片采用百万种彩色扫描，对黑白航片或卫片采用灰度扫描。
    - b、扫描分辨率的设置，根据扫描要求，对地形图的扫描一般采用300dpi或更高的分辨率。
    - c、针对一些特殊的需要，还可以调整亮度、对比度、色调、GAMMA曲线等。
    - d、设定扫描范围。
  - 扫描参数设置完后，即可通过扫描获得某个地区的栅格数据。
- 矢量化

# ArcGIS 下扫描矢量化

- 利用ArcCatalog创建必要的图层（shape file）
- 在ArcMap中利用Georeferencing进行图像校正
- 使用ArcScan进行矢量化
  - \ArcTutor\Editor\ExerciseData\Digitizing
  - \ArcTutor\ArcScan\

## 2.2 属性数据的采集

- 属性数据的录入主要采用键盘输入的方法，有时也可以辅助于字符识别软件。
- 当属性数据的数据量较小时，可以在输入几何数据的同时，用键盘输入；但当数据量较大时，一般与几何数据分别输入，并检查无误后转入到数据库中。

# §3 空间数据的坐标变换

- 几何纠正
  - 相似变换
  - 仿射变换
  - 二次多项式变换
- 投影变换
  - 解析变换法
  - 数值变换法
  - 数值解析变换法

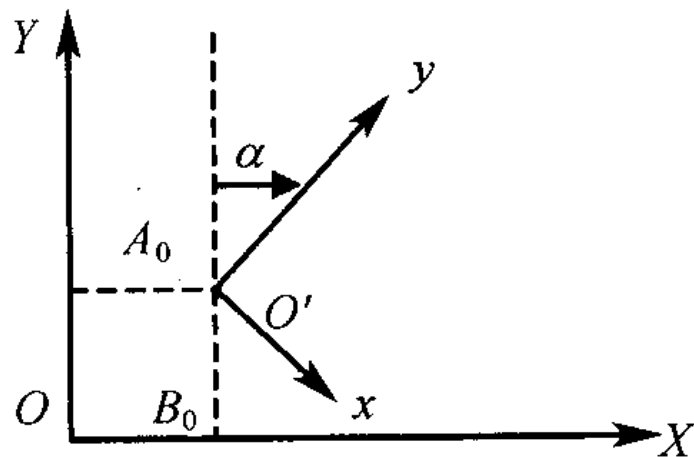
## 3.1 几何纠正

- 在图形编辑中，只能消除数字化产生的明显误差，而图纸变形产生的误差难以改正，因此要进行几何纠正。



# 相似变换

- 两坐标系之间的平移、旋转、缩放。



$$X = m(x \cos \alpha - y \sin \alpha) + A_0$$

$$Y = m(x \sin \alpha + y \cos \alpha) + B_0$$

令

$$A_1 = m \cos \alpha$$

$$B_1 = m \sin \alpha$$

则上式可简化为

$$X = A_0 + A_1 x - B_1 y$$

$$Y = B_0 + B_1 x + A_1 y$$

# 仿射变换

- 如果坐标在X、Y方向的比例因子不一致，或者说图纸存在仿射变形，此时需要采用仿射变换公式。令 $m_1$ 和 $m_2$ 分别表示X和Y方向的比例尺，则变换公式为：

$$X = (m_1 \cos \alpha)x - (m_1 \sin \alpha)y + A_0$$

- 令

$$Y = (m_2 \sin \alpha)x + (m_2 \cos \alpha)y + B_0$$

$$A_1 = m_1 \cos \alpha, \quad A_2 = -m_1 \sin \alpha$$

$$B_1 = m_2 \sin \alpha, \quad B_2 = m_2 \cos \alpha$$

- 则上式简化为

$$X = A_0 + A_1x + A_2y$$

$$Y = B_0 + B_1x + B_2y$$

# 二次多项式变换

- 二次变换适用于原图有非线性变形的情况，至少需要**6**对控制点的坐标及其理论值，才能求出待定系数。

$$\begin{cases} x' = a_0 + a_1x + a_2y + a_{11}x^2 + a_{12}xy + a_{22}y^2 \\ y' = b_0 + b_1x + b_2y + b_{11}x^2 + b_{12}xy + b_{22}y^2 \end{cases}$$

## 3.2 投影变换

- 一、地图投影的基本原理
- 地图投影就是指建立地球表面上的点与投影平面上点之间的一一对应关系。
- 二、地图投影的分类

- 变形分类（如图3-4）
  - 等角投影：投影前后角度不变
  - 等面积投影：投影前后面积不变；
  - 任意投影：角度、面积、长度均变形
- 投影面（如图3-5）
  - 横圆柱投影：投影面为横圆柱
  - 圆锥投影：投影面为圆锥
  - 方位投影：投影面为平面
- 投影面位置：
  - 正轴投影：投影面中心轴与地轴相互重合
  - 斜轴投影：投影面中心轴与地轴斜向相交
  - 横轴投影：投影面中心轴与地轴相互垂直  
  - 相切投影：投影面与椭球体相切
  - 相割投影：投影面与椭球体相割

- 常用的地图投影
- 高斯克吕格投影
- 墨卡托投影
- **UTM**投影
- 兰勃投影

- 四、地图投影的转换

- 当空间数据来自不同地图投影时，需要将一种投影的几何数据转换成所需投影的几何数据。

- 地图投影变换的实质是建立两平面场之间点的一一对应关系。

- 假定原图坐标为  $x, y$ ，新图坐标为  $X, Y$ ，则由旧坐标变换为新坐标的基本方程式为：

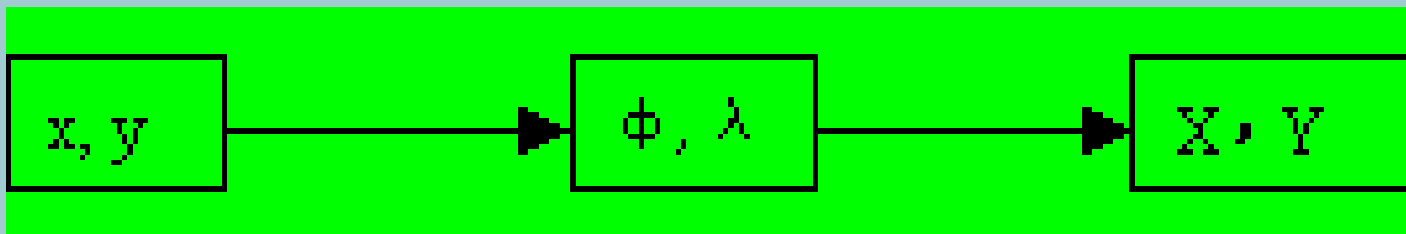
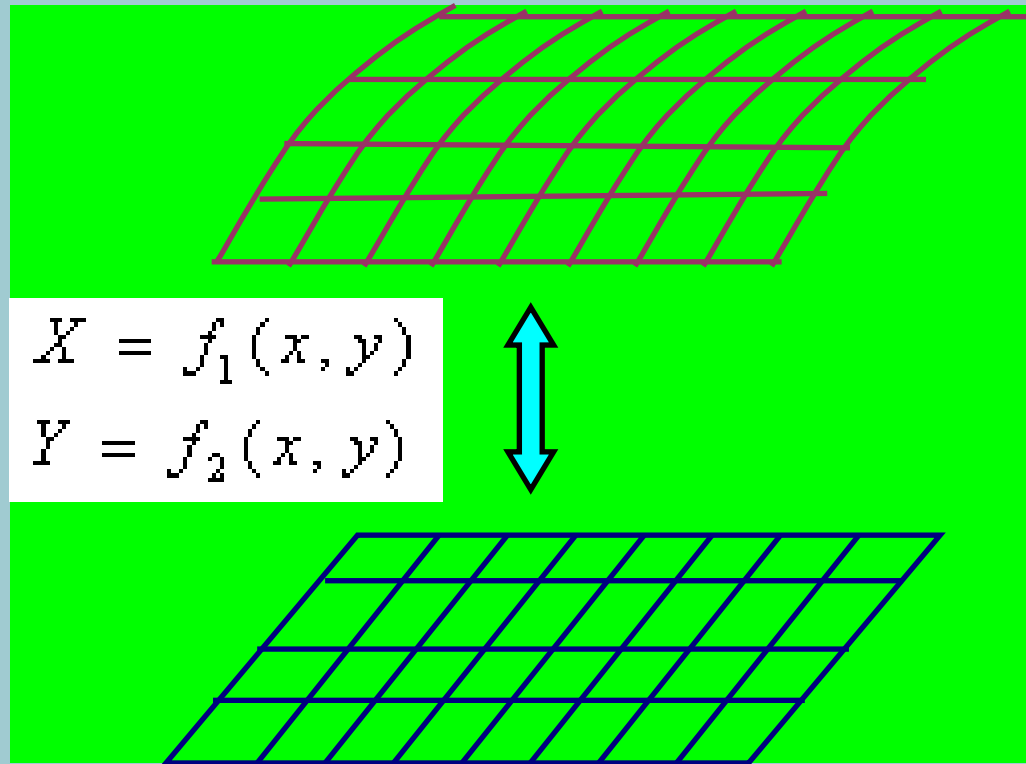
$$X = f_1(x, y)$$

$$Y = f_2(x, y)$$

- 投影变换方法：解析变换法、数值变换法、数值解析变换法

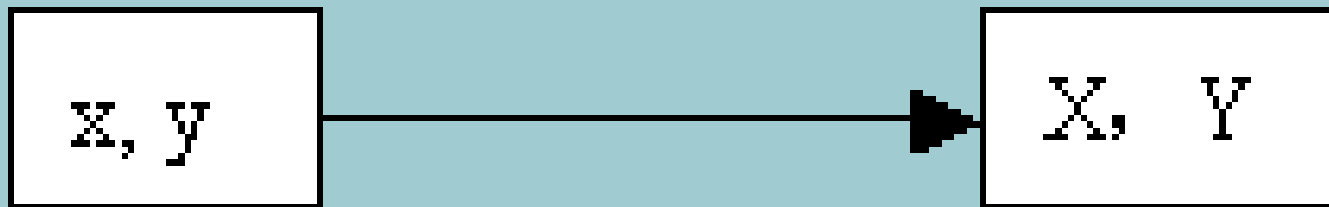
# 解析变换法

- 这类方法是找出两投影间坐标变换的解析计算公式。可分为反解变换法和正解变换法。
- 反解变换法(又称间接变换法): 先解出原地图投影点的地理坐标, 然后将其代入新图的投影公式中求得新坐标。即:



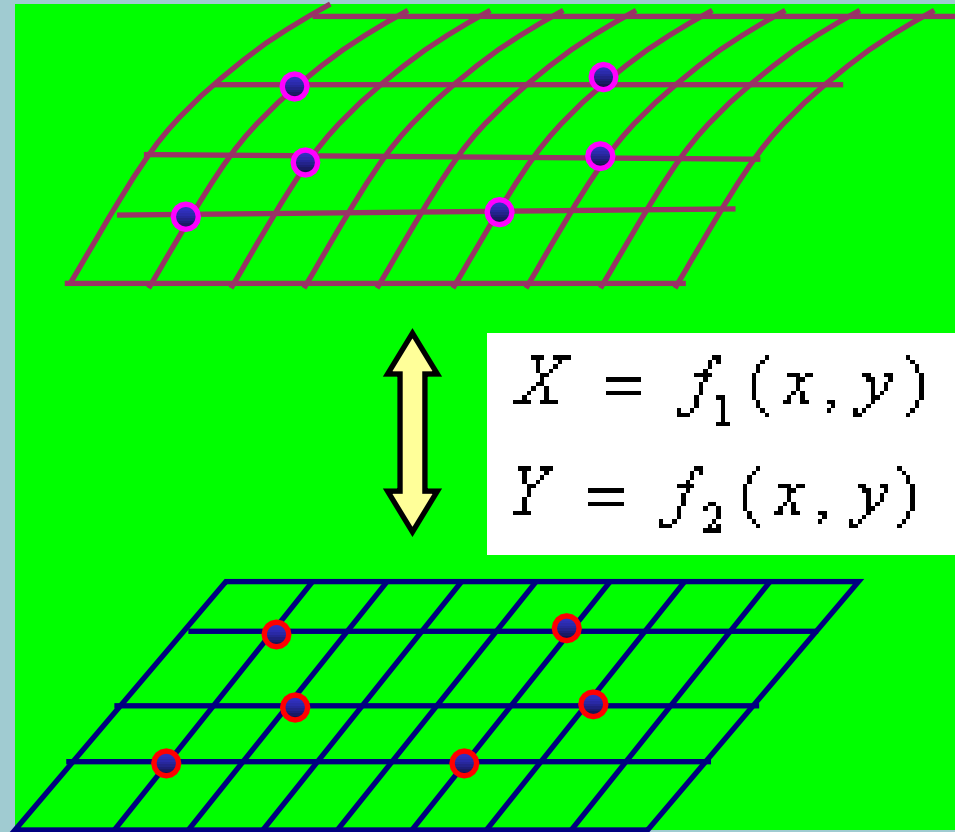


- 正解变换法(又称直接变换法)。这种方法不需要反解出原地图投影点的地理坐标的解析公式，而是直接求出两种投影点的直角坐标关系式。即：



# 数值变换法

- 如果原投影点的坐标解析式不知道，或不易求出两投影之间坐标的直接关系，可以采用多项式逼近的方法，即用数值变换法来建立两投影间的变换关系式。例如，可采用二元三次多项式进行变换。



$$\begin{cases} X = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 + a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3 \\ Y = b_{00} + b_{10}x + b_{01}y + b_{20}x^2 + b_{11}xy + b_{02}y^2 + b_{30}x^3 + b_{21}x^2y + b_{12}xy^2 + b_{03}y^3 \end{cases}$$

# 数值解析变换法

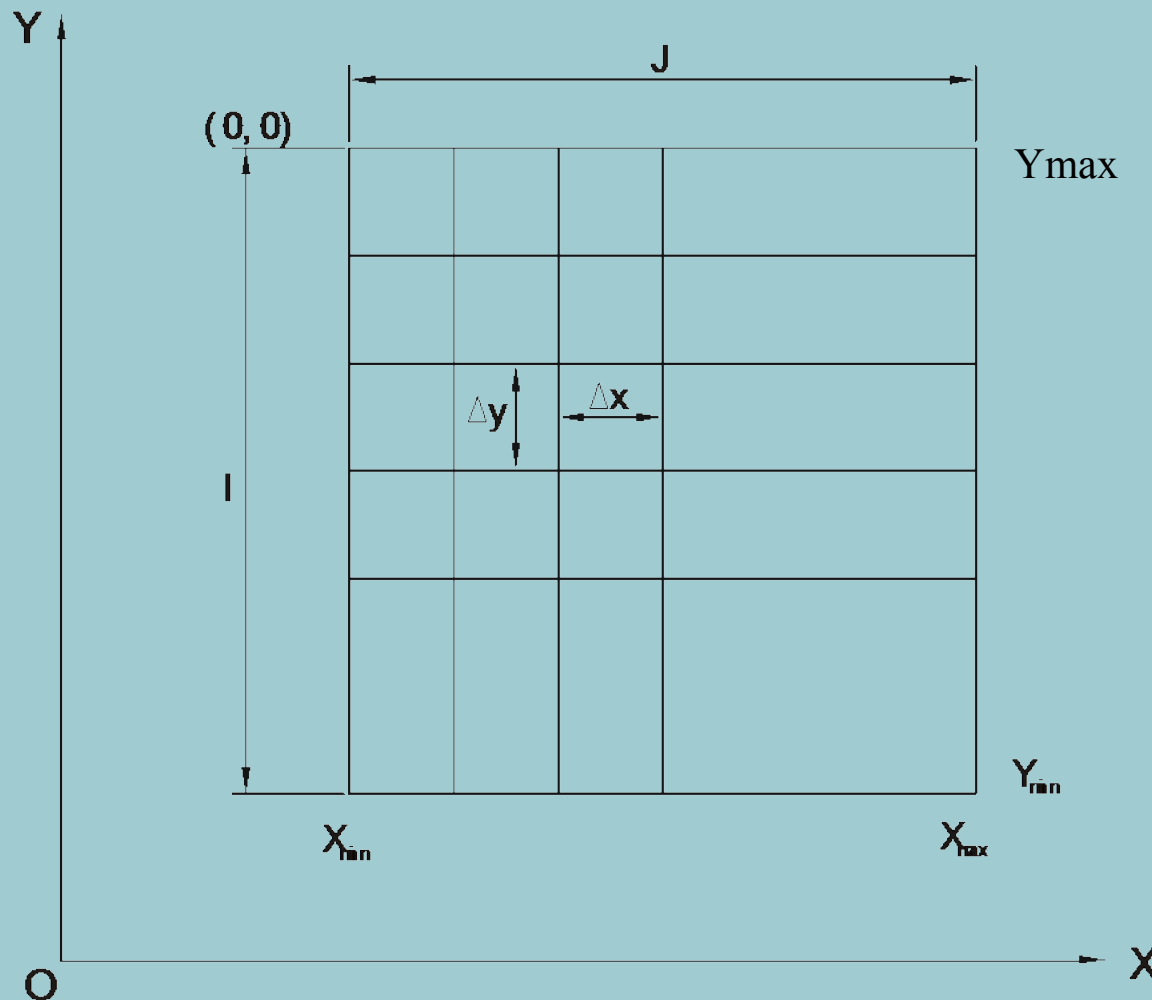
- 当已知新投影的公式，但不知原投影的公式时，可先通过数值变换求出原投影点的地理坐标 $\phi$ ， $\lambda$ ，然后代入新投影公式中，求出新投影点的坐标。即：



# §4空间数据结构的转换

# 4.1 矢量数据结构向栅格数据结构的转换

- 两种数据变换时，令直角坐标 $x$ 、 $y$ 分别与行和列平行。
- 由于矢量数据的基本要素是点、线、面，因而只要实现点、线、面的转换，就能实现整个线划图的转换。





# 确定栅格单元的大小

- 栅格单元的大小就是它的分辨率，应根据原图的精度，变换后的用途及存贮空间等因素予以决定。
  - 如果变换后要在一幅卫星图像匹配，最好采用与卫星图像相同的分辨率。
  - 如果作为地形分析用，地形起伏变化小时分辨率可以低些，栅格单元就可大些；而地形变化大时，分辨率就应当高些，栅格单元就要小些。

栅格单元的大小为 $\Delta x$ 和 $\Delta y$ ，  
设 $X_{\max}$ 、 $X_{\min}$ 和 $Y_{\max}$ 、 $Y_{\min}$ 分别表示全图X坐标和Y坐标的最大值与最小值， $N_I$ 、 $N_J$ 表示全图格网的行数和列数，它们之间的关系为

$$\left. \begin{aligned} \Delta X &= \frac{X_{\max} - X_{\min}}{N_J} \\ \Delta Y &= \frac{Y_{\max} - Y_{\min}}{N_I} \end{aligned} \right\}$$

# 点的变换

- 点 (X, Y) 的变换很简单，只要这个点落在某个栅格中，就属于那个栅格单元，其行、列号I、J可由下式求出：

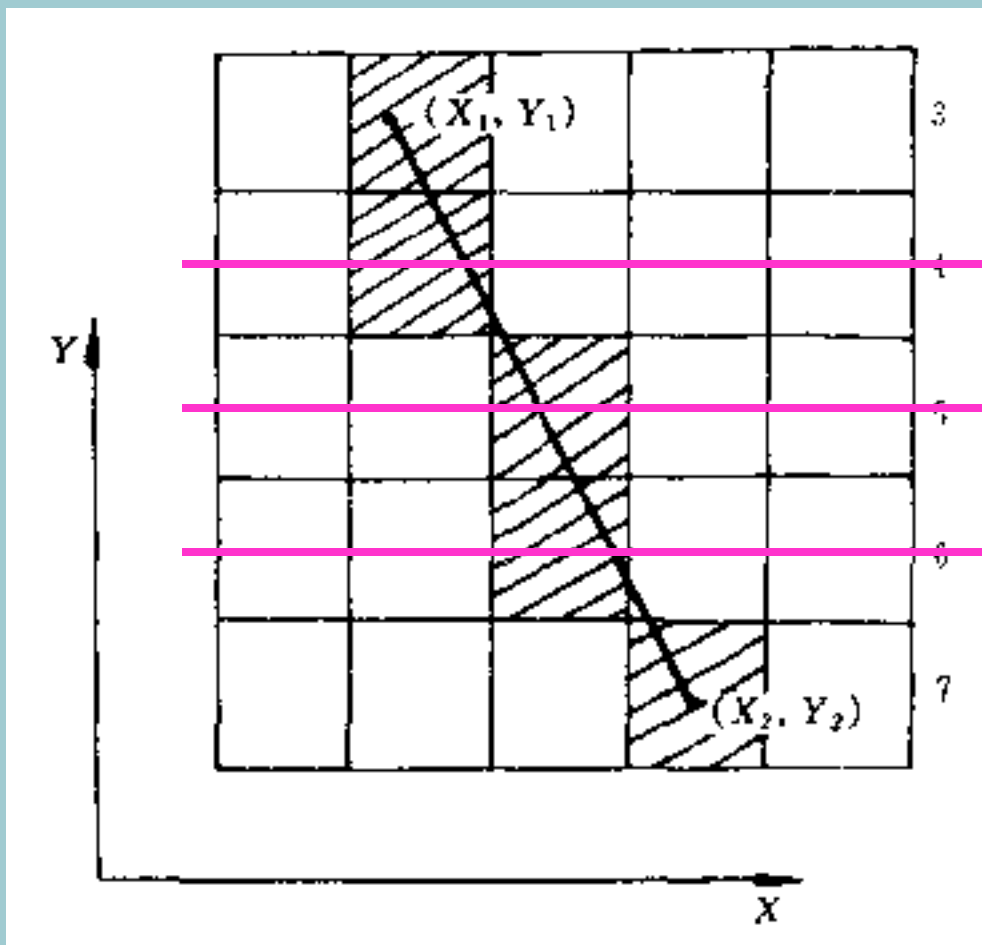
$$\left. \begin{aligned} I &= \text{INT} \left( \frac{Y_{\text{Max}} - Y}{\Delta Y} \right) \\ J &= \text{INT} \left( \frac{X - X_{\text{Min}}}{\Delta X} \right) \end{aligned} \right\}$$

- 式中INT表示取整函数。栅格点的值用点的属性表示。假定I, J从0开始起算。



# 线的变换

- 对于曲线可以近似地看成多个直线段组成的折线。每一条直线段需要求算它经过哪些格网单元。
- 设线段两端点的坐标为  $(X_1, Y_1)$ ,  $(X_2, Y_2)$ , 先求出这二个端点栅格单元的行列号, 然后求中间经过的栅格。



- 假设求出两端点的行号分别为n1、n2（如3和7），则中间网格的行号必为n1+1, n1+2, ……., n2（如4, 4, 6），其网格中心线的Y坐标应为

$$Y_i = Y_{\text{Max}} - \Delta_Y * \left(i + \frac{1}{2}\right)$$

- 它与直线段交点的X坐标为

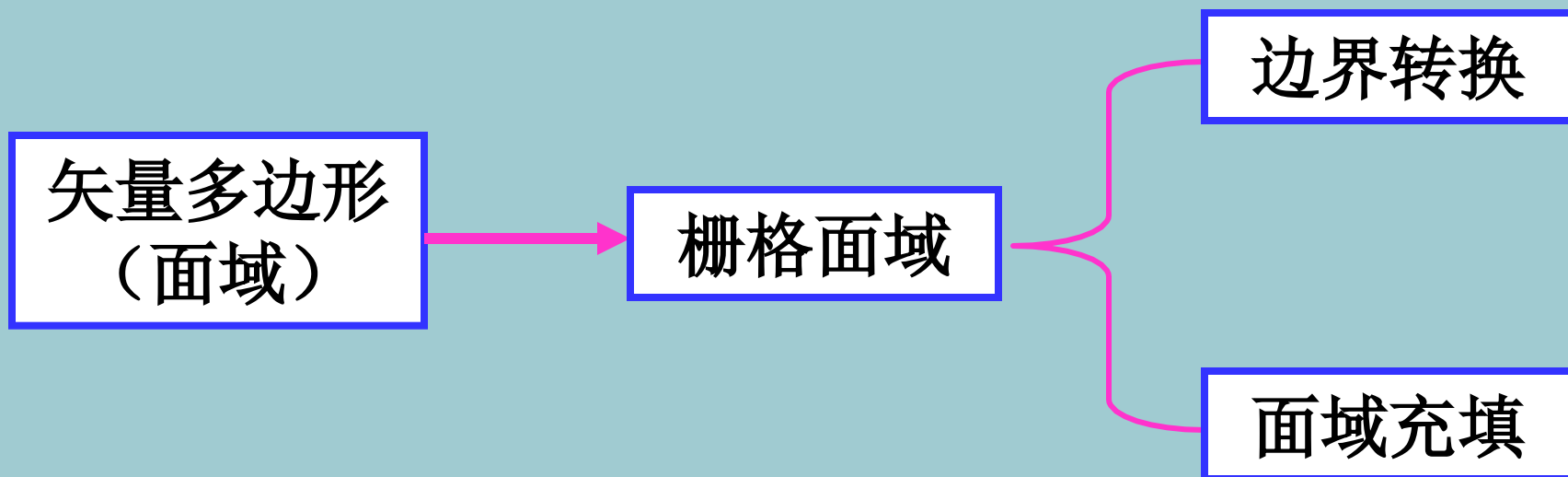
$$X_i = X_1 + \frac{X_2 - X_1}{Y_2 - Y_1} (Y_i - Y_1)$$

- 再由此Xi值可求出其列号J，即

$$J = \text{INT} \left( \frac{X_i - X_{\text{Min}}}{\Delta X} \right)$$

- 也可以先求出两端点的列号，再确定中间网格的行号。

# 面的充填



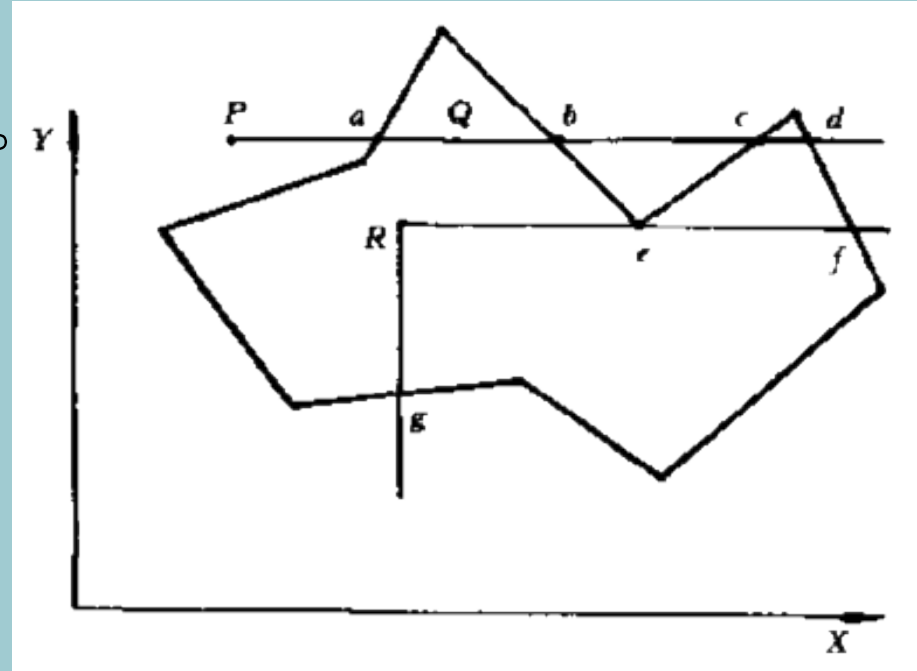
- 充填的关键问题是使计算机能正确判断哪些栅格单元在多边形之内，哪些栅格在多边形之外。

- 1.基于弧段数据的栅格的方法
- 基于弧段时针对拓扑结构的矢量数据进行栅格化。
- 步骤：按行或者按列作中心扫描线
- 对其中的扫描线,求出与所有矢量多边形的边界弧段的交点，记录其坐标，并用点的栅格化求出栅格坐标行列值，再根据弧段的左右多边形信息判断并记录交点左右多边形的数值。最后通过对一行所有交点按其坐标 $x$ 值从小到大排列，并参照左右多边形配对情况，逐段生成栅格数据。  
见图3-15

## 2. 多边形数据的栅格方法

### 1) 射线法

- 该法中常用的方法有平行线扫描法和铅垂线跌落法。前一种方法是从待检验的栅格单元作一平行于x轴的扫描线，当与多边形边界相交的点数为偶数时，则该栅格在多边形之外，当交点为奇数时，则该栅格在多边形之内。



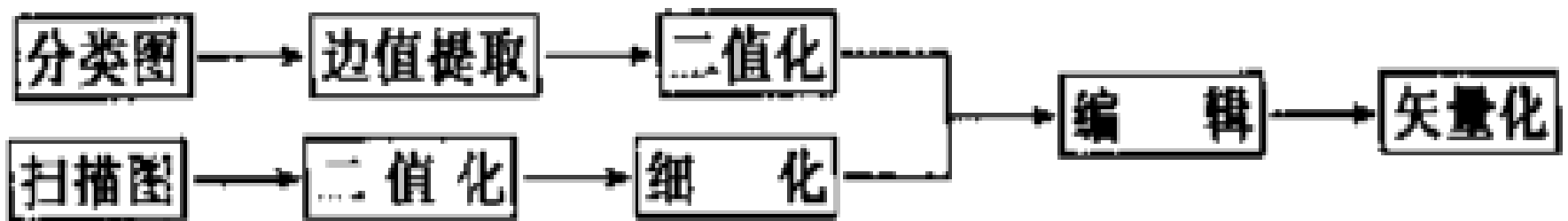
- 铅垂线跌落法则是从待检查的栅格作一垂直于X轴的直线，检查它与多边形边界交点的个数，奇数在多边形之内，而偶数在多边形之外。例如从R点作垂线只交多边形边界于x点，故R点在多边形之内。
- 为了避免误判，可以同时采用这两种方法检验，只要一种方法交点为奇数，该点就在多边形之内。遍历所有栅格单元，凡在多边形内的点均充填同一属性值。

## 2)扩散法

- 在多边形边界栅格确定后，寻找多边形中的一个栅格作为种子点，然后向其相邻的八个方向扩散。被扩散的栅格如果是边界栅格，就不再作为种子点向外扩散，否则就继续作为种子点向外扩散。重复上述过程直到所有种子点填满该多边形为止。

## 4.2 栅格数据结构向矢量数据结构的转换

- 栅格向矢量的转换过程比较复杂，它有两种情况：
  - 一种是本身为遥感影像或已栅格化的分类图
  - 另一种情况通常是从原来的线划图扫描得到的栅格图





# 基于图像数据的矢量化方法

## 边界提取

- 边界提取是遥感图像处理中的一个专门问题。
- 用一个 $2 \times 2$ 栅格的窗口，按顺序沿行列方向对栅格图像进行扫描。
  - 如果窗口内的四个网格点值相同，它们就属于一个等值区，而无边界通过，否则就存在多边形的边界或边界的结点。
  - 如果窗口内有两种栅格值，这四个栅格则均标识为边界点，同时保留原栅格的值。
  - 如果窗口内有三个以上不同的值，则标识为结点。

$a$	$a$
$b$	$b$

(1)

$a$	$b$
$a$	$b$

(2)

$a$	$a$
$a$	$b$

(3)

$a$	$a$
$b$	$a$

(4)

$a$	$b$
$b$	$b$

(5)

$a$	$b$
$a$	$a$

(6)

## 边界点的六种可能结构

## 结点的八种可能结构

$a$	$b$
$c$	$d$

(1)

$a$	$a$
$c$	$d$

(2)

$a$	$b$
$d$	$d$

(3)

$a$	$b$
$a$	$d$

(4)

$a$	$b$
$c$	$b$

(5)

$a$	$b$
$c$	$a$

(6)

$a$	$b$
$b$	$d$

(7)

$a$	$b$
$b$	$a$

(8)

# 二值化

- 所谓二值化就是将图像中的灰度取一个阈值，凡高于阈值的灰度取1，低于阈值的灰度取0。设阈值为  $\tau$ ，则二值化后的像元灰度值为

$$T(i, j) = \begin{cases} 1 & \text{当 } f(i, j) \geq \tau \text{ 时} \\ 0 & \text{当 } f(i, j) < \tau \text{ 时} \end{cases}$$

- 式中  $f(i, j)$  为原像元灰度。
- 二值图像中的图形用1表示，背景用0表示。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/408042105034007003>