

经济和社会  
事务

# 支持人口普查活动的 地理空间 基础结构 手册



联合国

经济和社会事务部  
统计司

方法研究

F辑 第103号

# 支持人口 普查活动的 地理空间 基础结构手册



联合国  
纽约，2010年

## 经济和社会事务部

联合国秘书处经济和社会事务部(经社部)在经济、社会和环境领域的全球政策与国家行动之间起着重要的桥梁作用。该部的工作主要涉及三个相互关联的领域：(一) 汇编、制作和分析范围广泛的经济、社会和环境数据与信息，供联合国会员国在审查共同问题和评价政策抉择时加以使用；(二) 促进会员国在许多政府间机构内就采取何种联合行动方针应对现有或新出现的全球挑战进行谈判；以及(三) 就将联合国各次会议和首脑会议制定的政策框架转化为国家一级方案的方式方法向有关政府提供咨询意见，并且通过技术援助协助国家能力建设。

### 说 明

本出版物中使用的名称以及材料的编写方式并不意味着联合国秘书处对任何国家、领土、城市或地区或其当局的法律地位，或者对其边界或界限的划分表示任何意见。

本出版物中使用的“国家”一词在适当情况下亦指领土或地区。

使用“发达地区”和“发展中地区”等名称是为了统计上的方便，未必表示对某一国家或地区在发展进程中所处发展阶段做出判断。

联合国文件都用英文大写字母附加数字编号，凡提到此种编号即指联合国某一文件。

ST/ESA/STAT/SER.F/103

联合国出版物  
出售品编号：C.09.XVII.8

ISBN: 978-92-1-661030-2

版权©联合国，2009年  
版权所有

由联合国印刷，纽约

## 前 言

联合国出版了《地理信息系统和数字化地图绘制手册》，用于2000年的一轮人口和住房普查。2000年的《手册》在人口普查地图绘制领域提供了有用的指导。不过，需要对该手册进行更新和审查，以考虑地理空间技术及其在统计实践，尤其是人口和住房普查方面应用的最新发展。

关于《2010年世界人口和住房普查方案》，统计委员会在其第三十六届会议上，请联合国统计司着手修订和更新《人口和住房普查的原则与建议》。委员会请统计司探讨一些相关的具体问题，包括地理信息系统(GIS)在收集和发布数据方面的应用。这些发展情况反映在2007年3月统计委员会通过的《人口和住房普查的原则与建议，修订2》中。

通过《原则与建议，修订2》后，联合国统计司开展了一系列活动，以促进和支持《2010年世界人口和住房普查方案》，并审查2000年《手册》，同时考虑到地理空间技术的最新进步。2007年5月，统计司在纽约组织了一个专家小组会议，讨论当前人口普查制图方面的实践和地理信息系统的应用。举办该会议旨在为修订2000年《手册》提供建议。统计司还在2007年第四季度和2008年年初举行了五个关于在人口普查制图活动中使用地理空间技术的讲习班。分别在非洲为讲英语的国家和讲法语的国家举办了两个讲习班，还为亚洲、加勒比和太平洋的岛屿地区各举办了一个讲习班。统计司聘请了一位顾问，David Rain，来协助编写《手册》修订本的草稿。2008年4月，在纽约组织了第二个专家小组会议，审查《手册》修订本的草稿。

新修订和命名的《支持人口普查活动的地理空间基础结构手册》反映了关于地理信息系统和人口普查制图的专家小组会议和区域讲习班提出的建议。这些机构强调指出，各国有必要将人口普查地理方案的使用作为一个连续的过程，而不只是一系列的制图和传播活动。还强调《手册》应表明当代地理空间技术和地理数据库的使用和应用如何有益于人口和住房普查的所有阶段。例如，《手册》应指出这些技术在普查的准备、点查、加工和发布阶段如何能够提高效率。

在这方面，重要的是《手册》应该向普查规划者和相关的工作人员提供一个技术指南，指出使他们能够更好地表达自己的需要并更有效地开展人口普查制图活动的当代方法、工具和最佳做法。简而言之，《手册》既涉及管理方面的需要，也涉及业务方面的需要，而且非常详尽。《手册》探讨了统计机构负责人和其他管理人员关心的组织问题和机构问题；清楚地阐明了人口普查制图者和普查员关心的技术问题和操作问题。

本《手册》分为六章和七个附件。第一章是简要介绍，第二章涉及统计机构负责人关心的管理问题，在重组国家统计局办公室，以便充分利用地理空间基础结构中应该考虑这些问题。第三章提供了技术方面的内容，供数据处理主管或地图绘制/地理信息系统负责人使用，以便实际建立点查区一级的数字化地理数据库。第四章继续把重点放在技术方面，阐述了建立点查区地理数据库的过程以及利用全球定位系统和遥感等技术进步来进行修正，必要时与实地工作相结合。第五章探讨了制作点查所需的地图的过程，讨论了关于地理数据库的论述中没有谈到的业务重点。第六章涉及使用地理空间基础结构来发布普查结果。附件一至七为规划和执行普查项目的地理空间解决办法提供了便捷的参考。

在修订过程中，联合国秘书处咨询了代表世界各个地区的人口普查绘图人员和地理信息系统专家，以便审查和最终完成《手册》。《手册》还包含一些专家提供的各国在将地理信息系统、全球定位系统和数字化地图绘制技术用于人口普查方面的实践例子。《手册》是联合国统计司的一名顾问，David Rain，在统计司团队的协助下编写的。

# 目 录

	页 次
前言 .....	iii
章 节	
<b>第一章 引言</b> .....	<b>1</b>
A. 《手册》前言和基本原理 .....	1
B. 《手册》的范围、目标和概要 .....	3
C. 逐章概述 .....	3
<b>第二章 国家统计局办公室和其他决策者需要考虑的管理事项</b> .....	<b>7</b>
A. 引言 .....	7
B. 地图在人口普查中的作用 .....	12
C. 从地图到地理数据库：制图方面的革命在继续 .....	12
D. 对分类统计数据的需求越来越大 .....	14
E. 投资于地理空间技术：成本和收益 .....	16
F. 在国家统计局办公室实施地理空间项目的关键成功要素 .....	18
G. 利用地理空间工具规划人口普查工作 .....	19
H. 需求评估和确定地理选择 .....	19
1. 用户需求评估 .....	19
2. 确定产品 .....	21
3. 地理数据的选择 .....	21
4. 人力资源和能力建设 .....	22
I. 机构合作：国家空间数据基础结构：确保与其他政府部门的兼容 .....	25
J. 标准 .....	29
K. 合作 .....	29
L. 概要和结论 .....	30
<b>第三章 为人口普查建立一个点查区级别的数据库</b> .....	<b>31</b>
A. 引言 .....	31
B. 国家人口普查地理学的定义 .....	32
1. 行政结构 .....	32
2. 行政管理与其他统计汇报或管理单位之间的关系 .....	33
3. 点查区地面划界的标准和程序 .....	35
4. 管理(人口普查小组领导)区划界 .....	37
5. 点查区的地理编码 .....	37

	页次
6. 人口普查数据库的组成 . . . . .	38
7. 点查区的设计与过去的人口普查保持一致 . . . . .	39
C. 用于点查区划界的地理数据源 . . . . .	41
1. 所需地图的类型 . . . . .	41
2. 现有资源库存 . . . . .	43
3. 导入现有的数字化数据 . . . . .	44
4. 地理数据的转换：从模拟到数字 . . . . .	47
(a) 扫描 . . . . .	49
(i) 其他考虑 . . . . .	51
(ii) 扫描的利弊 . . . . .	53
(b) 数字化处理 . . . . .	53
数字化的利弊 . . . . .	55
(c) 编辑 . . . . .	56
(d) 建立和维持拓扑结构 . . . . .	57
(e) 数字地图的集成 . . . . .	57
(f) 制定地理基准 . . . . .	57
(g) 投影和数据的变化 . . . . .	59
(h) 分立地图块的集成 . . . . .	59
D. 运行一个点查区数据库 . . . . .	61
1. 关系型数据库 . . . . .	61
各种关系型数据库和地理数据库结构 . . . . .	62
2. 数据库内容(数据建模)的定义 . . . . .	64
E. 数据质量问题 . . . . .	64
1. 准确度要求 . . . . .	64
2. 质量控制 . . . . .	67
3. 将国家领土划分为操作单元 . . . . .	68
4. 数字化的行政底图 . . . . .	68
5. 处理不相交的地区单元 . . . . .	69
6. 面积计算 . . . . .	69
F. 元数据的开发 . . . . .	71
G. 概要和结论 . . . . .	73
<b>第四章 利用全球定位系统和遥感数据整合实地工作 . . . . .</b>	<b>75</b>
A. 全球定位系统(GPS) . . . . .	76
1. 全球定位系统的工作方式 . . . . .	76
2. 全球定位系统的准确度 . . . . .	77
3. 差分全球定位系统 . . . . .	77
4. 其他全球卫星导航系统 . . . . .	79

	页 次
5. 人口普查制图应用领域使用的全球定位系统 . . . . .	80
6. 一些与全球定位系统有关的特殊的制图任务 . . . . .	81
7. 对使用全球定位系统的培训要求 . . . . .	82
8. 总结：全球定位系统的优缺点 . . . . .	83
B. 利用便携式计算机的集成的实地制图系统 . . . . .	84
C. 卫星遥感 . . . . .	86
1. 利用图像来实地核查在人口普查总部制作的点查区地图 . . . . .	86
2. 卫星遥感的原理 . . . . .	87
3. 遥感数据的分辨率 . . . . .	88
4. 卫星遥感数据在线资源 . . . . .	90
5. 将遥感数据用于人口分析 . . . . .	91
6. 卫星遥感数据的优点和缺点 . . . . .	93
D. 航空摄影 . . . . .	93
1. 航空摄影概述 . . . . .	93
2. 航摄照片在人口普查制图方面的应用 . . . . .	96
3. 航空摄影的实施与机构问题 . . . . .	96
4. 航摄照片的优缺点 . . . . .	97
E. 概要和结论 . . . . .	98
<b>第五章 在人口普查期间使用地理数据库(地图) . . . . .</b>	<b>99</b>
A. 引言：人口普查的点查期间使用地理空间工具 . . . . .	99
B. 质量保证、点查区地图制作和数据库维护 . . . . .	99
1. 概述 . . . . .	99
2. 草图制作和质量保证程序 . . . . .	100
3. 边界线和属性文件的匹配以及总图的打印 . . . . .	100
4. 质量保证 . . . . .	101
5. 地方当局的核查和最终的行政单位检查 . . . . .	102
6. 点查区地图的制作(包括地图打印) . . . . .	103
C. 在人口普查的点查期间使用地理空间基础结构 . . . . .	107
1. 运用数字化地图作人口普查的后勤保障 . . . . .	107
2. 监督人口普查工作的运作过程 . . . . .	108
3. 点查员在人口普查期间利用地图的准则 . . . . .	109
4. 在点查期间更新和修正点查区地图 . . . . .	109
D. 概要和结论 . . . . .	109
<b>第六章 发布人口普查结果、产品和服务的地理数据库 . . . . .</b>	<b>111</b>
A. 引言 . . . . .	111
B. 点查结束后以及两次人口普查之间的工作 . . . . .	112



	页次
直接工作 . . . . .	112
1. 整理点查员所做的更新和改动. . . . .	112
2. 将收集单位与制表单位或统计单位进行整合. . . . .	112
3. 数据库的存档. . . . .	114
4. 数据库的维护：继续进行制图的好处 . . . . .	114
C. 发布地理人口普查产品 . . . . .	115
1. 规划数据的发布 . . . . .	115
2. 披露和数据隐私性考虑：区差问题. . . . .	117
3. 地理人口普查产品和服务的市场开发 . . . . .	119
4. 外联和教育 . . . . .	119
5. 可能的产品清单 . . . . .	120
(a) 等效文件与兼容文件 . . . . .	120
(b) 基准地图库 . . . . .	120
(c) 地名字典和矩心文件 . . . . .	120
6. 用于出版目的的专题地图 . . . . .	121
(a) 地图的威力 . . . . .	121
(b) 人口普查数据专题地图 . . . . .	122
(c) 数字化人口普查地图册 . . . . .	124
(i) 静态地图册 . . . . .	124
(ii) 动态地图册 . . . . .	126
(d) 空间分析技术. . . . .	128
(e) 地图制作和公布问题：输出的类型. . . . .	131
(f) 绘图工具和软件 . . . . .	132
7. 输出方式：数字化文件 . . . . .	132
(a) 数据格式. . . . .	133
(b) 光栅图像格式. . . . .	134
(c) 地理信息系统数据格式 . . . . .	136
(d) 坐标数据. . . . .	137
(e) 表列数据. . . . .	138
D. 打印 . . . . .	138
1. 概述. . . . .	138
2. 打印机类型 . . . . .	139
3. 商业化印刷 . . . . .	140
E. 用于发布目的的数字地理数据 . . . . .	141
1. 针对潜在用户的数字化数据发布策略 . . . . .	142
(a) 数据内容的定义 . . . . .	143
(i) 发布何种层次的数据? . . . . .	143

	页 次
(ii) 是一个大型地理信息系统数据库，还是一系列 人口普查数据库? . . . . .	143
(iii) 如何确定边界与数据库的集成密集度? . . . . .	144
(iv) 需要提供多少元数据? . . . . .	144
(b) 文件命名约定 . . . . .	144
(c) 压缩 . . . . .	145
(d) 文档，包括数据字典 . . . . .	145
(e) 可交付的数据产品的质量控制和保证 . . . . .	146
2. 法律和商业问题 . . . . .	147
(a) 数据版权 . . . . .	147
(b) 地理数据商业化的权衡 . . . . .	148
(c) 责任问题 . . . . .	149
3. 通过因特网制图 . . . . .	150
(a) 侧重服务器的模式 . . . . .	151
(b) 侧重客户的模式 . . . . .	152
(c) 混合模式 . . . . .	153
(d) 人口普查数据发布面临的机遇，包括MapServer. . . . .	153
F. 概要和结论 . . . . .	156
书目和参考资料 . . . . .	157
 <b>附 件</b>	
一、地理信息系统 . . . . .	161
A. 地理信息系统概述 . . . . .	161
B. 地理信息系统数据模式 . . . . .	165
C. 地理信息系统的能力 . . . . .	170
二、坐标系统和地图投影 . . . . .	175
A. 引言 . . . . .	175
B. 坐标 . . . . .	175
C. 地图投影的特性 . . . . .	179
D. 更为精确的制图方法：地理基准法 . . . . .	180
E. 地图比例尺 . . . . .	183
F. 地理基准实例 . . . . .	185
G. 实际考虑 . . . . .	187
三、数据模型操作 . . . . .	189
A. 引言 . . . . .	189
B. 关键术语的定义 . . . . .	189
C. 模板实例 . . . . .	189

	页次
四、用于发布的数据字典实例 . . . . .	193
五、专题地图设计 . . . . .	197
A. 引言 . . . . .	197
B. 地图设计原则 . . . . .	197
C. 数据分类 . . . . .	217
D. 色彩选择 . . . . .	227
E. 地图图例的设计 . . . . .	229
F. 说明问题的地图 . . . . .	230
六、专业术语 . . . . .	235
七、有用的地址和URLs. . . . .	255
<b>方    框</b>	
II.1. 四个国家的案例研究 . . . . .	9
II.2. 技术和成本障碍 . . . . .	17
II.3. 数据共享合作的三个例子 . . . . .	26
II.4. 国际机构的参与和协调 . . . . .	28
III.1 地理空间软件的选择标准：COTS、图像分析和FOSS选择 . . . . .	44
IV.1. 全球定位系统方面的经验案例研究：斐济 . . . . .	85
IV.2 数字化正射影像地图的开发 . . . . .	95
V.1. 印度人口普查实地地图的制作 . . . . .	110
VI.1. 为发布人口普查数据进行网络绘图的案例研究：加拿大 . . . . .	155
<b>图</b>	
II.1. 人口普查地理工作规划中的各个阶段 . . . . .	20
III.1. 一个普通的人口普查地理结构 . . . . .	33
III.2. 嵌套式行政结构图示 . . . . .	34
III.3. 一个普通的点查区编码方案 . . . . .	38
III.4. 数字化空间普查数据库的组成 . . . . .	39
III.5. 人口普查地理数据库开发的各个阶段 . . . . .	42
III.6. 数据转换过程的相互制约 . . . . .	47
III.7. 进给式扫描仪的照片 . . . . .	48
III.8. 半自动向量化处理 . . . . .	51
III.9. 扫描图像数据的向量化和光滑化 . . . . .	52
III.10. 数码化表 . . . . .	54
III.11. 平视数字化 . . . . .	55
III.12. 几种常见的数字化错误 . . . . .	56

	页 次
III.13. 平移、缩放、旋转 . . . . .	58
III.14. 以数字化单位表示的地图；以真实世界坐标表示的地图 . . . . .	58
III.15. 连接相邻的数字化地图块 . . . . .	60
III.16. 连接相邻地图块后的边缘配接 . . . . .	60
III.17. 实体表实例——点查区 . . . . .	62
III.18. 关系型数据表 . . . . .	63
III.19. 逻辑准确度 . . . . .	65
III.20. 不保持定位准确度将产生的问题 . . . . .	66
III.21. 处理由若干多边形组成的行政单位 . . . . .	70
III.22. 占若干行政单位一大部分面积的湖泊 . . . . .	70
IV.1. 差分全球定位系统 . . . . .	78
IV.2. 在全色卫星图像上划分点查区边界 . . . . .	86
IV.3. 遥感过程 . . . . .	87
IV.4. 电磁波谱 . . . . .	87
IV.5. 航摄照片和卫星图像像素大小图解 . . . . .	88
IV.6. 摄影胶片与扫描图像的对比 . . . . .	94
IV.7. 地形所造成的失真 . . . . .	95
V.1. 质量保证、输出成品和数据库维护阶段 . . . . .	100
V.2. 数字化点查区地图的组成例子 . . . . .	105
V.3. 城市点查区地图的例子 . . . . .	106
VI.1. 人口普查制表和报告单位的例子 . . . . .	114
VI.2. 统计披露中的区差问题 . . . . .	118
VI.3. 静态数字化人口普查地图册的不同设计展示方法 . . . . .	125
VI.4. 乌克兰动态地图集中的一个屏幕截图 . . . . .	127
VI.5. 一个多边形对象的缓冲 . . . . .	129
VI.6. 创建轮廓线的直线内插法例子 . . . . .	130
VI.7. 关于泰森多边形的说明 . . . . .	130
VI.8. 统计地图的例子 . . . . .	131
VI.9. 数字化印刷过程 . . . . .	141
VI.10. 因特网制图：侧重服务器的模式 . . . . .	151
VI.11. 因特网制图：侧重客户的模式 . . . . .	152
A1.1. 地理信息系统的基础 . . . . .	162
A1.2. 存储在地理信息系统内的各种类型的信息 . . . . .	163
A1.3. 数据层——作为索引系统的空间 . . . . .	164
A1.4. 点、线和多边形 . . . . .	165

	页次
A1.5. 向量数据模式：面条模式与拓扑模式的对照 . . . . .	166
A1.6. 向量型地理信息系统中存储的空间和非空间数据 . . . . .	167
A1.7. 光栅数据文件实例 . . . . .	168
A1.8. 向量和光栅数据模式均可用于显示离散和连续数据 . . . . .	169
A2.1. 平面坐标系和极坐标系 . . . . .	176
A2.2. 球面坐标：纬度/经度基准系统 . . . . .	176
A2.3. 地图投影过程(方位投影)图解 . . . . .	177
A2.4. 各种地图投影法 . . . . .	178
A2.5. 构成投影的各种方式 . . . . .	179
A2.6. 常用的地图投影法 . . . . .	181
A2.7. 圆球与椭圆球的对比 . . . . .	182
A2.8. 通用横向墨卡托基准体系 . . . . .	183
A2.9. 联合国总部大厦在UTM基准体系中的位置 . . . . .	184
A2.10. 地图上的控制点 . . . . .	185
A5.1. 通过选择颜色或灰度级构成视觉上的层次 . . . . .	198
A5.2. 行政单位和主要城市中心地图实例 . . . . .	201
A5.3. 人口密度专题地图实例 . . . . .	202
A5.4. 在显示空间特征时泛化的效果 . . . . .	202
A5.5. 为了便于地理信息系统进行表达，复杂的现实状况有时需要简化 . . . . .	203
A5.6. 变量的测量 . . . . .	203
A5.7. 表示多边形、线和点的图形变量 . . . . .	204
A5.8. 点状密度地图 . . . . .	206
A5.9. 点密度图与等值线图的组合 . . . . .	207
A5.10. 绘制离散点对象图 . . . . .	208
A5.11. 用于点状和面积特征物的比例符号 . . . . .	208
A5.12. 图画式符号与简单图形符号的对比 . . . . .	209
A5.13. 用简单图形符号表示流动的规模 and 方向 . . . . .	210
A5.14. 通过不同的地图符号数量表示每一种特征的数据值 . . . . .	210
A5.15. 饼图式地图 . . . . .	211
A5.16. 等值线与饼图结合的地图 . . . . .	212
A5.17. 通过直方图显示不同时期变化的地图 . . . . .	212
A5.18. 地图与人口金字塔的组合 . . . . .	213
A5.19. 在地图上显示人口的性别比率 . . . . .	213
A5.20. 各种表示地区间移民流量的方法 . . . . .	215
A5.21. 移民迁入和迁出的表示法 . . . . .	215

	页 次
A5.22. 各种表示连续性数据的地图绘制方法. . . . .	216
A5.23. 相等的间距. . . . .	219
A5.24. 分位点(等频)制图 . . . . .	221
A5.25. 标准偏差 . . . . .	223
A5.26. 根据标准偏差指定等级的灰度 . . . . .	224
A5.27. 自然中断点. . . . .	225
A5.28. 灰度地图内不同类型的图例 . . . . .	230
A5.29. 用图例显示统计数据的分布 . . . . .	231
A5.30. 将实色和填充线符号组合在同一地图上表现两个变量 . . . . .	231
A5.31. 与双向表格等价的地图 . . . . .	232
A5.32. 小型多幅地图——说明随时间产生的变化. . . . .	234

## 表

IV.1. 空间分辨率非常高的民用卫星遥感产品 . . . . .	90
VI.1. 新旧点查区单位的对比 . . . . .	113
VI.2. 可用于普查地图册的专题地图清单 . . . . .	123
A2.1. 纽约联合国总部大厦按不同基准椭圆球投影后得出的坐标 . . . . .	182
A2.2. 参数转换 . . . . .	187
A3.1. 用于定义空间数据模型的汇编信息 . . . . .	190
A3.2. 实例：具有三级管理的国家行政单位. . . . .	191
A5.1. 不同分类技术的评估 . . . . .	227
A5.2. 灰度级和颜色的选择 . . . . .	229



# 第一章 引言

## A. 《手册》前言和基本原理

1.1. 《支持人口普查活动的地理空间基础结构手册》力求承认和利用其前身，即2000年出版的《地理信息系统和数字化地图绘制手册》的成就。

1.2. 技术方面的主要进步包括普遍可以获得个人电脑、手提电脑、全球定位系统、地理信息系统软件和低成本的航摄图像和卫星图像。这些进步使国家统计局办公室获得了新工具，可以收集更准确、及时和公正的人口信息。新技术的出现确实是新《手册》背后的驱动力。与此同时，要认识到一点，那就是采纳这类新方法将对国家统计局办公室的领导能力构成挑战，导致国家统计局办公室的组织发生变化。执行新的地理空间功能贯穿人口普查方案的始终。

1.3. 本《手册》认为，对国家统计局办公室来说，一些最大的挑战不只是技术方面的，还是组织、制度和管理方面的。大多数会员国已经开始利用合适于自己的方案需要的地理信息科学和技术。《手册》的目的之一就是满足这些需要。在许多国家进行国家人口普查的法律或宪法授权仍然有效。与此同时，变化包括更多地利用人口普查数据来进行灾害管理和其他许多方面(更完整的清单见第二章)。其他变化是能够获得更多的各种可用格式的数据；以及新技术，这使得数据的收集、分析和存储比以往更加容易。

1.4. 在为这些用途制作数据中，各国发现它们可以通过所谓的国家空间数据基础结构(NSDI)，发挥其他政府机构的优势。<sup>1</sup> 国家空间数据基础结构是一种制度安排，可以在不同级别的政府中进行数据共享和合作，包括国家、区域和地方各级。其他制度问题，例如筹资、人员配备和项目管理等基本问题，虽然本身不是技术问题，但对地理空间人口普查项目的成功将产生影响。“一次建造，多次使用”的原则适用于国家人口普查地理空间数据库的建立，因为该数据库一旦建成，可以在不同的国家背景下用于多种目的。

1.5. 本《手册》为围绕数字化地理基准数据库重新安排人口普查地图绘制和分析任务提供了建设性的办法。围绕这样的“地理数据库”建立统计机构的结构要求进行周密的规划，不仅是因为地理信息系统需要进行前期投资，而且还因为需要为能力建设开展工作，以便分析人口普查数据并及时向公众发布产品。作为一个产业，地理信息系统的成功建立在地理空间信息能够解决问题和支持决策这一能力之上。

<sup>1</sup> 国家空间数据基础结构是获取、处理、储存、分配和改进地理空间数据的利用所必需的技术、政策、标准和人力资源的组合。国家空间数据基础结构的概念部分包括(a) 一个制度框架，确定对基本数据集建立、维持和执行标准的政策、法律和行政支持；(b) 定义基本数据集的技术特征的标准；(c) 需要大地测量框架、地形和地籍数据库的基本数据集；(d) 使用户能够确定和获得访问基本数据集的途径的技术框架(见GSDI Cookbook, 2000)。



1.6. 发展地理信息系统能力可能意味着改变国家统计局办公室的传统组织结构，尤其是扩大现有的“地图绘制单位”，使其成为一个更大的和有更多功能的地理核心，能够满足人口普查地图绘制方面的所有需要。完成这一重组可能需要开展持续的有资金支持的活动，配备有专门知识的人员，以帮助国家统计局办公室完成自己十年一轮的任务。一个专用的地理核心将需要一个团队，这个团队不仅要有地理空间技能，而且还要有使人口普查工作现代化的业务目标。训练有素的工作人员将遵循严格的时间表，以确保及时向实地的点查员提供经过修正的详细地图，以开展人口普查工作。一个数字化点查区数据库将包含各地的人口估计数以及点查区边界，为了表示小区域在地理空间上界定了这些边界。任务将需要持续的管理，包括实地工作(在向工作人员提供越野车方面可能花费巨大)与遥感数据收集相结合，从而使人口普查更准确和更具成本效益。

1.7. 规划地理方案的运作不会太快。由于实施复杂的组织重组需要很长的时间，规划工作需要提前数年进行。就点查而言，点查员和小组负责人需要十分详细的大比例地图。在足够恰当的范围内制作详细的地理数据，与地理信息系统的其他数据层一起使用，并且发布单位由点查区或人口聚居区组成，这意味着国家统计局办公室方面要做大量规划工作。利用一个全面的关系型数据库系统来建立一个专用的地理核心或数据中心，这可能需要进行组织方面的重组。基本的数据层，例如国家统计局办公室确定的人口统计和行政边界，可以在许多用户中共享，避免重复劳动的费用。这两个数据层是任何一个国家空间数据基础结构的基本组成部分。事实上，就国家空间数据基础结构方面的工作而言，国家统计局办公室的贡献可能是：

- (a) 一个空间地理数据库，带国家点查区(即人口普查期间为点查划分的领土单元)的多边形和属性信息。一个普通的数据库可以对农业和人口普查提供帮助，正如最近的国家经验所揭示的。可以在点查区一级发布人口普查数据，或对新的小区域发布单位，例如人口聚集区，计算总数；
- (b) 一个数字化行政边界库，从省一级到市一级(甚至还有地块一级)。如果国家统计局办公室对边界进行了全面的测量并进行了实地修正，那么可以避免为其他目的(如重划选区)再次进行领土测量，从而节省大量费用；
- (c) 一个国家地名字典，包括人口居住地的地理名称和坐标(在联合国人道主义界被称为“P-编码”)。与人口估计数组合在一起时，人道主义组织可以将其用于发展和紧急应对目的。可以以类似的方式使用关于住宅单位的其他向量格式的数据。

1.8. 为了在其他组织中共享地理信息系统格式的地理数据和人口数据，国家统计局办公室应遵循国家一级制定的地理基准工作和元数据的标准。对于用于安排人口普查地域的行政分类制度，需要做特别的考虑。本《手册》将把这项工作称为“地理编码”，因为它是人口信息与该人口在地球表面的位置的联系点。

1.9. 最重要的是,《手册》强调了有必要制定务实的计划,以便利用地理信息系统和其他地理空间技术的力量来使人口普查工作现代化,并获得更好的结果和分析。还强调了有必要将人口普查制图的范围扩大到一国的国家空间框架。

## B. 《手册》的范围、目标和概要

1.10. 撰写本《手册》的主要推动力来自近来数字制图技术和地理分析技术的迅速发展以及对小地域地理基准人口数据需求的不断增长。任何一个着手开展普查项目的国家都需要进行周密的规划以便尽可能降低成本并从所需的地理空间活动中得到最大好处。除了给国家统计局内部地理活动的关键支持者提供信息,本《手册》旨在提供有关技术和方法的资料,对任何特定国家选择合适的工具和程序提供支持。

1.11. 由于有大量选择存在、各国的条件和可用资源又极其不同,因此每一种情况下的选择显然也不会相同。因此本《手册》被设计为一份实用参考文件,一本“食谱”,举例说明了地理空间技术在人口普查工作各个阶段的作用。各国必须根据自己的人口普查方案和国家规划,从各种选择中选取适合自己情况的地图绘制方法。为每一种情况确定最佳技术和方法的组合取决于以下因素:本国已有的地理资源、技术资源和人员、可用资金以及完成人口普查地理任务的时间框架。

1.12. 本《手册》并非一本通用的地理信息系统手册。它不包含运行特殊软件所需的命令和程序,也不针对特殊程序提出建议。它也不是用于进行人口普查的一本综合指南。本《手册》尊重普查制图的优良传统,认为很多国家成功使用了几个世纪的传统模拟地图绘制技术并未过时。有关这一主题的主要参考依据之一——《人口普查与调查的制图》(美国人口普查局,1978年)无论对新手还是资深绘图人员仍是有用的资料,尤其是关于制图方案的组织和控制、点查区的划界以及统计区的章节。本《手册》大量使用了前人的资料。然而,随着技术的发展,出现了许多完成普查制图任务的更好方法。因此,本《手册》通过提供有关最新技术的信息,旨在补充以前的指导方针。

1.13. 本《手册》面向两个主要的受众群体,一个是关心投资于地理空间技术的成本和收益的管理人员群体,另一个是负责执行人口普查计划的地理方面的技术人员群体。《手册》的五个章节让人对地理信息系统和制图概念有了一个基本了解。对于不太熟悉这些主题的读者,附件一和二简要概述了两个主题。特别是,地图投影和坐标系统在利用地理信息系统的项目中比在传统的基于草图的方法中是一个更为重要的主题。

## C. 逐章概述

1.14. 本《手册》第二至六章涵盖人口普查的点查活动之前、之中和之后地理空间技术可以做出贡献的领域。经过修订的《手册》编排方式反映了许多会员国为促使向数字化人口普查业务转变而做出的业务决定。

1. 15. 第二章涵盖统计机构负责人要考虑的管理问题，提出了在重组国家统计办公室以便充分利用地理空间基础结构时要考虑的一些问题。规划应满足对分类统计数据的新需要，包括灾害管理数据，要求进行反向工作，即从发布新产品到一路下来在各个步骤所需的必要条件。论述了效率方面的论据，包括成本效益分析，目的是强调地理空间技术作为一项战略投资的作用。接下来探讨了对于执行地理空间工作来说关键的成功因素，然后论述了详细的规划过程。这包括进行需求评估；确定产品和可选方案；人员配备以及人力资源方面的其他考虑；以及通过国家空间数据基础结构与其他数据制作者和用户的机构间合作。

1. 16. 第三章为数据处理主管或制图/地理信息系统负责人提供了技术方面的内容，以便实际建立一个数字化地理数据库。该章节一个步骤一个步骤地阐述了建立一个点查区级别的数据库的过程，首先是地理编码、行政结构以及人口普查数据库各个组成部分的重要性。接下来谈到了点查区划界的问题，然后论述了用于点查区划界的现有地理数据的来源，包括数字的和模拟的来源。随后谈到了通过扫描和数字化得到的数字输入，接着论述了地图集成、拓扑结构和地理基准工作。建立一个地理数据库可能会引起与准确度和精确度以及质量保证程序有关的新的数据质量问题。该章节在最后详细论述了元数据技术，强调了文件记录的重要性，并为内部和外部的数据用户提出了策略方面的建议。

1. 17. 第四章仍然把重点放在技术方面，阐述了建立点查区一级的地理数据库的过程，以及如何利用技术的进步，例如全球卫星定位系统和遥感技术来进行修正，并且在必要时将这些技术与实地工作相结合。这里使用的是一种分类方法，为的是把重点放在自上次人口普查以来发生变化的区域。阐述了全球卫星定位系统的基础，以及差分全球卫星定位系统的技术新进步、培训要求以及对手提电脑的使用。接下来是关于卫星遥感技术的部分，列出了可以利用的遥感来源，以及可视化工具，例如谷歌地球 (Google Earth)，接下来的部分谈到了将航空摄影用于人口普查工作。

1. 18. 第五章谈的是绘制点查所需地图的过程，包括论述地理数据库时没有谈到的业务重点。该章节包括关于质量保证、实地使用的地图的绘制、地图要素和设计、地图打印和分发以及将地图用于人口普查后勤基础工作的部分。它强调指出项目管理方法必须包括遇到挫折时的应急计划。

1. 19. 第六章谈的是利用地理空间基础结构来发布人口普查结果。对于人口普查后的阶段，该章节运用了与前面章节相似的方案，但利用了地理数据库(对特殊地点进行了地理编码)的概念，作为发布人口普查产品的中心数据界面。谈到的主题包括将收集数据的单位并入发布数据的单位、地理数据库的维护和归档，以及计划人口普查产品。关于空间分析的部分提供了形象化地解释人口普查数据的例子和可选办法。讨论的其他问题包括披露和保密问题(关系到发布小区域数据)以及国家统计机构可以选择通过出售有附加值的产品，如

CD-ROM或DVD数据，来增加收入。还探讨了地理数据产品，如地图阅读器、用于商业地理信息系统软件包或在线产品的基于属性的空间文件。

1.20. 附件为那些规划和执行人口普查项目地理空间解决方案的人提供了方便的参考。附件一提供了关于地理信息系统数据模型以及准确度和精确度级别的基本信息。附件二介绍了坐标系和地图投影，特别关注将数据输入地理信息系统的问题。附件三涉及数据模型的建立，包括一个作为例子的模板。附件四提供了一个数据字典的例子。附件五对用于人口普查数据分析和显示的专题地图设计进行了一些指点，包括常规的等值线地图。附件六包含一个常用的地理信息系统词汇的术语表。附件七含有一个清单，列出了联络信息和有用的网址，供收集更多信息之用。



## 第二章

### 国家统计局办公室和其他决策者 需要考虑的管理事项

#### A. 引言

2.1. 本章面向的是国家统计局办公室主管、人口普查主任和地理部门负责人。本章主要讨论制度问题(也就是说,非地理信息系统的技术内容),重点是与使用地理空间技术有关的各种考虑事项。国家统计局办公室可以在较短的时间内制作出准确度较高的地理基准数据,但只有对活动进行了周密规划才行。本章包含国家经验的例子,说明了地理空间技术对人口普查工作的用处。

2.2. 对利用地理空间功能的人口普查进行规划可能需要对国家统计局办公室进行重组,把地理——包括地图绘制和地理信息系统以及诸如地理编码和行政边界等主题——置于核心位置。由于没有一种技术解决方案能够使每个国家统计局办公室都以同样的方式使其普查活动现代化,本《手册》提供了一系列可以使国家统计局办公室在适当的技术和经验级别采纳新技术的选择。它把这些选择设想为一个调色板,国家统计局办公室可以从中选择,以完成特殊任务,可以根据国家和国家统计局办公室的大小全部或部分地调整和采纳这些任务。在这一过程的全过程,重要的是强调必须早日开始规划工作,并且在必要时请求别人的帮助。

2.3. 在人口普查时要按照国家政府的要求计算准确的人口数,对于从科学的角度了解人口,人口普查也是社会最重要的手段之一。地理空间技术改变了为社会利益和促进社会、经济或可持续发展展示信息的方式。利用地理的核心组织原则,地理空间技术可以在几乎任何行业大大提高生产率。通过有序地显示以地点为中心的数据可以从地理的角度安排信息,有效地将社会观察与观察的地点联系起来。作为一项数据收集工作,一个以地理为基准的人口普查活动为国家了解详细的人口信息提供了绝好的机会。

2.4. 一般说来,在政府工作中,地理信息系统和地理空间技术,包括遥感卫星图像和全球定位系统,其作用正在发生变化,各国越来越清楚看到围绕一个暗含地理的空间模型组织政府信息的好处。当政府机构利用地理空间工具交流结果,进行查询时,使用地理数据库前景广阔,并且在许多情况下,在诸如效率和客户服务等重要领域让人获益匪浅。

2.5. 对“电子政务”进行的许多研究证明了采纳技术带来的好处。O'Looney(2002年)强调了政府与商界、民间团体和公民个人之间的界限日益模糊。利用信息技术来提供服务使得互动水平提高,并且较少的工作人员就

可以做到。Garson(2003年)强调了与之相伴的问责水平的提高,尤其是当信息的使用者能够检查计算并质疑结果时。Khosrow-Pour(2005年)探讨了需要高层更多参与的新的系统管理挑战;“24/7政府”对产品和服务的提供有明显的影 响,这意味着公民/客户需要提高自助服务的能力。White(2007年)强调指出有必要制定透明的关于因特网使用的内部政策,还敦促高层决策者更好地利用承包商来实现长期影响较小的近期的人力资源目标。

2.6. 由于软件、硬件和重组的初始投资很大,在人口普查活动中使用地理空间技术需要进行周密计划。地理信息系统的行业领导者,联合王国的David Rhind承认,地理信息系统不只是一种技术,还日益成为商界、政府和学术界——从某种意义上说全都像企业一样运营——运转方式的一部分。像任何战略投资一样,必须仔细权衡地理信息的成本和收益。加拿大的Roger Tomlinson坚持认为周密的规划会导致成功,不论是对于地理信息系统还是任何其他技术的实施来说都是如此。

2.7. 人口普查是一次精心组织的全国性活动,将在未来数年对政策产生影响。国家统计局负责人和其他主管应对此进行规划,想一想国家统计局办公室在接下来的几年内将提供什么产品。通过从发行到规划和数据需要的反向工作,他们应该决定需要发布哪些常规的人口普查产品——包括表格、总数、年龄和性别分类数字。他们还力图规划新产品,例如地图册、带分类数据或小区域数据或微观数据的DVD,以及可以满足大量新数据用户需要的电子地图,从而提高总体的客户满意度。利用地理学的力量,这类新功能尽在掌握之中。正如Tomlinson所说的,地理信息系统是一个特别横向的技术,广泛应用于工业和知识方面。应采纳这一技术来满足特殊需要。应该考虑国家统计局的战略目的。主要目标很可能是使人口普查既准确,又准时,而且在预算之内。特殊的部级目标或者组织所承担的任务应该明确,这些目标和任务将如何影响人口普查计划也应该明确。Tomlinson要求各组织考虑如何利用全球定位系统和地理数据库等新技术来改善其战略目标。他建议在人口普查的具体规划期间对技术的采纳进行广泛的成本效益分析。要求主管们在与用户讨论并拟定随后可用于满足要求的信息产品说明书后,规划自己机构的产品。他们随后应利用已经确定的要求进行数据设计;为规划的数据选择一种符合逻辑的数据模型并确定系统要求;最后对计划的实施做出规划。

2.8. 在人口普查方面,国家统计局办公室应计划重组处理和发布系统,以交付客户要求的产品。在组织结构方面可以感受到潜在的广泛影响,首先是围绕组织内部一个存储和提供数据的数据中心或信息核心选择中心化业务。

2.9. “以地理为中心的人口普查”意味着围绕地理组织人口普查工作。对于许多数字能力强大的国家统计局办公室来说,要进行投资的地方常常是纸张式的点查区地图从模拟向数字转变,这意味着细致的扫描和修正程序,以便点查区可以被用作新的数字化地理数据库的基础。这些新的地理数据库随后可以与航摄照片或卫星图像等遥感图像进行对比,然后利用全球定位系统进行实地校正。采纳扫描、卫星图像以及全球定位系统等技术使得地图绘制单位可以把

工作的重点放在自上次人口普查以来最需要更新的地区。

2.10. 在点查期间，地理空间技术有助于对后勤进行规划并让决策者了解最新进展情况。在完成计数后，这些技术可以通过信息发布使信息更多地进入社会，尤其是利用因特网。“地理空间一站式”方式有助于协调向数以千计的新数据用户提供人口普查产品，并使现有用户能够申请新的产品和服务。

2.11. 对于任何一项新技术来说，一个关键问题是能力建设。许多国家统计局办公室主任不认为他们拥有预算或机构能力来重组其机构。这应该更多地被看作是一个远景规划的问题，而不是预算问题。主管们应该提前五年，甚至十年进行思考；国家统计局需要的信息产品远比现在能够提供的更为详细。

#### 方框II.1

#### 四个国家的案例研究

##### 1. 纳米比亚

纳米比亚在筹备2001年的人口普查中开始了它的数字化地图绘制和地理信息系统方案，其主要目的是有效生产实地工作所需的底图。在一个咨询公司的帮助下建立了地理信息系统基础结构，包括利用全球定位系统收集和数字化的人口普查所需的空间数据。纳米比亚的土地面积约为824 000平方公里，人口180万，因此覆盖整个领土在后勤方面是一个挑战。中央统计局对13个地区和107个选区划分了点查区的边界，并且利用由九个数字组成的惟一标识符对其进行地理编码。还创建了其他边界层，包括国家公园、农场、镇、市和地方。一些挑战包括：借以开展工作的组织结构当时没有包括地理信息系统，缺乏训练有素的工作人员，缺乏培训，最初没有空间数据，边界问题和无法进入某些地区。为了解决人力资源缺乏的问题，纳米比亚与工学院建立了伙伴关系，开设地理信息系统和信息与通信技术方面的专门课程。为了2011年的人口普查，纳米比亚将利用使用户能够创建自己的地图的Postgress开放源软件建立一个以网络为基础的地理信息系统。将利用Oracle集中储存和读取数据。中央统计局正在计划收集居住单元，并利用这些信息划分点查区的边界(如果想了解更多信息，请联系Ottillie Mwazi: omwazi@npc.gov.na)。

##### 2. 不丹

不丹在2005年利用地理信息系统进行了人口和住房普查。国家统计局在划定6 800个点查区的边界时保障了完全覆盖，不丹国土面积47 000平方公里，总人口约230万。在将住房列入清单的工作中，所有结构，包括庇护所、临时住房和岩洞，都利用全球定位系统进行了定位。从比例为1:50 000的地形图上提取了空间信息，在地方行政单位设立了点查区。对于数据的发布，国家统计局建立了一个网络服务器，提供一套行政边界。发布的指标地图包括就业、健康、住房、供水、能源以及卫生。计划要求利用数字底层来规划下一次不丹的人口普查(如果想了解更多信息，请联系Thinley Jyamtshow Wangdi: thinly\_j@yahoo.com)。

##### 3. 圣卢西亚

为进行农业普查，圣卢西亚在1995年开始建设自己在地理信息系统方面的能力。中央统计局意识到过去使用的手绘的等高线测量图不适合于为人口普查确定住房单元的位置这一任务。通过土地和测量局、规划部内部的自然规划股、林业局和农业部，以及私营部门方面Cable和Wireless以及St. Lucia Electricity Services的协助，中央统计局对点查区地图进行了数字化，并且利用高端全球定位系统来从实地收集空间数据。2004年，在自然规划部测绘局的协助下，中央统计局能够利用岛屿的航摄照片了。中央统计局设计了定居点边界，对每一个居住单元使用经纬度，从而可以界定圣卢西亚国内的任何区域。利用九位数字对定居点进行了地理编码。在准备转入数字化地图绘制活动期间，圣卢西



亚遇到了人力资源方面的挑战,包括人员的摩擦。尽管如此,它还是能够向现有的工作人员提供地理信息系统方面的培训并完成自己的指定任务(如果想了解更多信息,请联系Sherma Lawrence: sherma\_l@slucia.com)。

#### 4. 巴西

2007年,巴西地理和统计研究所进行了一次普查活动,包括三方面的工作:农业普查、人口计数和出于统计目的的国家地址登记。这三方面的工作是以综合的方式同时进行的。支持这一活动的领土数据库包括一套地图和记录文件(地籍),其中可以看到巴西领土被划分成了小的地理区域或“点查区”。

领土数据库本身总共包含了249 068个点查区。其中,在2007年的普查活动期间视察了162 770个点查区,包括70 085个农村地区和92 685个城市地区。

巴西地理和统计研究所所有27个州办公室,其总部位于各州的首府。州办公室协调530个地方办公室的活动。每个地方办公室管辖一组自治市。这样一来,所有5 564个巴西自治市都能包括在普查活动中。每个州办公室的总部有一个领土数据库部门,负责相关州的人口普查地图绘制工作。此外,在六个州办公室有测绘管理股,在全国各地区开展工作,为领土数据库部门制作自治市的数字化(统计)地图提供支持。

这就是支持2007年点查区地图绘制的基础结构。在530个地理和统计研究所的地方办公室更新了现有的纸张式地图,领土数据库部门在前面提到的GGC的支持下进行了数字化处理。所有这一切都由巴西地理和统计研究所地球科学局下属的两个管理单位集中协调,这两个管理单位分别负责城市和农村的人口普查地图绘制活动。在这一阶段后,地图被送到各自治市的人口普查委员会进行修订和核准,人口普查委员会由相关的地方政府和社会的代表组成。

为了确保在全国范围内统一进行领土数据库的更新工作,使用了计算机网络自动化系统,以及详细的业务手册。组织了区域和地方培训方案。技术协调员开展了监督活动。通过一个伴随和控制生产的系统,按照既定的时间表对项目的进展进行了监督。

为2007年的人口普查活动建立领土数据库是一项十分艰巨的工作。这项劳动密集型工作需要在地籍和地形两方面全面清查现有地图以及其他附属文件;与第三方建立合作伙伴关系以便交换文件和合并领土信息;对地图文件进行更新和数字化;制成地方、自治市和地籍地图;编制点查区数字化边界文件;以及绘制个人数码助理所用的地图并支持公布人口普查结果。

为自治市/农村地区制作点查区地图依靠的是巴西地理和统计研究所以及巴西军队地理事务局制作和拥有的系统的地形图。根据这些被用作输入数据的地形图,通过半自动城市地图绘制系统(SisCart)进行了制作,该系统是在Bentley/Intergraph's MicroStation/MGE图表平台和微软的文字数字数据管理产品Access97的基础上在Visual Basic 6.0中专门为巴西地理和统计研究所开发的。

SisCart极大地方便了市级地图的绘制。以分散方式运行的SisCart提供的功能支持各种关键任务,包括地图投影和比例的均一化;包含市级地图的地形图的地理编码;相邻地图的边缘匹配;根据自治市的周界对地图边缘进行修整;以及编辑框架和脚注数据。

城市地区的点查区地图制作依赖于地籍图,地籍图的比例范围为1:2 000至1:10 000。这些地图是由州和地方政府机构、公用事业公司以及其他机构制作的。用于城市地图制作的系统依据的是MicroStation平台(完全定制以满足巴西地理和统计研究所的要求)。该系统可以进行从各种各样的来源和坐标系提取地图数据所需的一切转化。它还可以对地图更新任务提供支持,从实地工作和办公室活动中取得输入数据。

在巴西2007年人口普查活动中引入的主要创新之一是以统计为目的的国家地址登记，该登记簿最初是根据2000年人口普查点查区的纪录编制的，随后通过2007年的人口普查实地工作进行了更新。2007年引入的技术创新，尤其是使用与全球定位系统集成的PDA，使得巴西地理和统计研究所制作了新的领土数据库产品，例如：

- PDF格式的70 085个农村点查区和92 685个城市点查区地图。
- PDF格式在农村和城市点查区说明。
- JPG格式的农村点查区地图。
- 70 085张Google Earth地理编码图像。
- 向量格式的市/点查区数字化边界，包括城市周界和所有27个联邦单位中的孤立的地区，约77 000个多边形。

所有与领土数据库有关的资料都到了地方办公室和收集站里的点查员手中，地方办公室和收集站是实地数据收集活动开始和结束中的集中地点。通过PDA以数字化的方式提供了2007年人口普查点查区的地图并且被点查员记录在笔记本上，连同点查区的周界说明。

历史上，处理人口普查收集而来的如此大量的信息是领土数据库部门关注的一个问题。随着PDA的采纳，要显示在地图上的实地数据可以被转化并储存在巴西地理和统计研究所的服务器上，以便进行进一步处理、分析，并纳入新地图的发布，从而减轻领土数据库部门的工作量。在收集活动期间，点查员指出了打印出来的地图上需修订的地方，另外，通过分析SisCart要添加的更新数据的相关性，点查员也对修订工作做出了贡献。

下表详细列出了2007年人口普查期间领土数据库更新工作的操作成本，包括其在人口普查活动总预算中的份额。

根据招聘联邦政府工作人员的立法，外包事务办公室组织了公共采购，进行临时工作人员的招聘。2004年，为实施几个阶段的领土数据库更新工作，招聘了500名测绘助理人员，他们在巴西地理和统计研究所领土数据库部门的地方办公室以及DGC总部工作。

2007年为开展数据收集工作聘请了大约68 000名点查员和18 000名人口普查监督员。所有点查员都接受了实地培训，上过有关领土数据库的设计、各点查区的相关特征以及领土数据库修订程序的课程（如果想了解更多信息，请联系：Rafael Castaneda: rafael.march@ibge.gov.br）。

领土数据库更新工作的成本					
年份	数据库更新		2007年普查		数据库更新%
	(R\$)		US\$ <sup>a</sup>		
2004	10 774 885.19	—	3 468 273.47	—	
2006	4 041 240.44	179 200 904.00	1 867 917.93	82 829 167.55	
2007	834 380.58	428 919 454.00	433 354.41	222 769 011.11	
2008	—	1 457 000.00	—	835 292.09	
合计	15 650 506.21	609 577 358.00	5 769 545.81	306 433 470.76	1.88%

<sup>a</sup> 参考值：每年6月30日巴西雷阿尔(R\$)对美元的汇率：2008年\$1.7443、2007年\$1.9254、2006年\$2.1635以及2004年\$3.10670。

## B. 地图在人口普查中的作用

2. 12. 几个世纪以来，人类一直用地图表示环境。地图用以标示地点、距离、方向和地域面积。地图还用来展示地理关系、差异、群集和格局。地图也用于导航、探险、说明以及公共部门和私营部门中的通信。几乎所有科学探索领域都需要以这种或那种形式使用地图。总之，对于许多专业或学术研究领域而言，地图都是一件不可缺少的工具。

2. 13. 长期以来，制图一直是人口普查工作中不可缺少的组成部分。地图在人口普查中的传统功能一直是以图形形式支持点查，以及合并人口普查结果。从最近的几次人口普查活动来看，点查很少有不依赖详尽的地图支持完成的。

2. 14. 一般来说，制图在人口普查过程中可以达到以下几个目的：

- (a) 地图确保覆盖面，便利人口普查活动(点查前)。人口普查办公室必须确保把全国的每一户和每个人都统计在内，并且不能重复统计任何一户和任何人。为此，人口普查地理学家要把全国分割成小的数据采集单位。显示点查区的地图因而提供了一个重要的控制手段，可以保证人口普查的覆盖面；
- (b) 通过地图支持数据收集并帮助监督各项普查活动(在点查阶段)。在人口普查过程中，地图能够确保点查人员很容易确定出他们的指定地理区域，在该区域中，他们将对住户进行点查。人口普查负责人还可通过绘制好的地图制定计划和控制各项任务。因此，地图在监督人口普查活动的进展方面也将发挥作用。普查工作负责人可以借助地图进行战略规划，分配工作，确定有问题的区域并很快采取补救行动；
- (c) 地图使人口普查结果更易于展示、分析和发布(点查结束后)。人口普查结果的图形展示是一种非常有效的手段，有利于人口普查结果的形象化说明。它有助于体现重要的人口统计指数和社会指数的地域特征。因此，地图是公共部门和私营部门进行政策分析的一个不可缺少的部分。

## C. 从地图到地理数据库：制图方面的革命在继续

2. 15. 今天，地图只是包含在广义的地理信息中的一种信息显示形式，并且这种地理信息最常采用的形式是地理基准数据库(或地理数据库)。比起其他领域，绘图学受信息革命的影响要晚一些。早期的计算机擅长于存储数字和文本。相对而言，地图更复杂，而且数字化制图需要大量数据存储容量和快速计算能力。另外，制图基本上属于图形应用，而早期计算机的图形输出容量有限。因此，1960年代最早使用计算机完成的制图，范围只限于几个国家的政府和学术项目。商品化的地理信息系统直到1980年代才在能力上达到了一定水

平，并导致在下述领域得到迅速应用：地方和地区政府、城市规划、环境部门、探矿、公用事业部门和商业市场以及房地产公司等。这对国家统计局办公室开始使用制图技术产生了影响。

2.16. 新的信息来源还缩短了从项目设计到可操作数据库建立完成的时间。最重要的新的近期发展体现在导航、遥感、图像分析、数据处理和因特网绘图方面。在从调查到环境监测乃至交通调度等一系列领域，大面积数据采集工作由于全球定位系统的使用而发生了突破性的革新。新一代的商业高分辨率卫星可以拍摄到地球表面上几乎任何地方的照片。这些照片显示的详尽信息足以支持大量的制图应用，包括人口普查。由于全球定位系统技术与遥感数据采集的紧密集成，高精度数字化制图的费用已显著下降。

2.17. 地理信息系统极大地受益于计算机科学各领域的发展。比较先进的数据库软件可以处理大量的数字化地图信息。人口普查用户为储存、检索和显示地理对象开发了数据模型。先进的视频技术使我们能够越来越复杂地展现我们的环境。地理信息系统数据显示功能不仅远远超过了二维静态显示效果，而且可以提供三维动态模型。正如文本信息的输入可以通过光学字符识别技术来完成那样，快速而高分辨率的扫描技术和复杂的软件也可以加速地图数据的转换，而这种地图数据的转换工作过去只能依赖人工进行数字化处理。这种赋能技术使得数据制作者能够从纸张式地图库向数字化地图转变，并更新现有信息。不过，仍然需要知识来增加后来通过采纳技术成为可能的产品和服务的价值。

2.18. 类似的进步发生在地理数据发布领域。所有主要的地理信息系统销售商现在都提供可以通过因特网获取地理空间数据库或地理数据库的工具。各级政府部门也使用这一技术向公众廉价而快速地提供大量空间信息。因特网正在作为最重要的数据发布手段，代替印制的地图和数字化媒体。相应地，地图本身不再是静态的物品，而是不断变化的地理数据库的动态快照。

2.19. 因特网的制图程序说明，数字化空间信息工具的使用正在变得越来越便宜和便于使用。虽然高端地理信息系统软件包仍然需要大量的培训才能掌握，但是桌面制图软件已经不比使用标准的业务软件更复杂。数字化制图在标准计算机应用方面也正在紧密集成，例如，电子数据表格、图形和商业管理软件等。

2.20. 一些统计机构是地理信息系统的早期用户。人口、社会和经济统计数字是公共规划和管理的基础。社会经济指数的空间分布指导着地区发展、服务提供和许多其他方面的政策制定。数字化技术可以实现更完善的管理、更快的检索和数据显示的改善。因此，在地理和统计之间一直存在着紧密的联系——例如，在拉丁美洲的许多国家里，国家统计局和测绘局设在同一幢办公楼里，这一事实本身就说明了这种联系。由于紧密集成的地理信息系统可降低采集、编辑和发布数据的成本，又可缩短时间，因此在统计方面的应用给国家统计局办公室带来很大的益处。地理信息系统使统计机构可以提供更多种类的服务，因此，极大地增加了在数据采集方面的投资回报。

2.21. 制图自动化、地理信息系统以及其他地理空间工具使得可以更有效地制作点查区地图和反映人口普查结果的专题地图。另外，技术的进步和利用新的数据来源的地理信息系统的新任务，例如遥感和全球定位地理信息系统启用的位置记录，扩大了国家统计局办公室中的地理表现能力。不过，应根据人口普查机构负责人确定的战略目标审慎使用新技术。新技术不是一颗能够解决所有组织问题的魔法弹，而是必须仔细规划。

#### D. 对分类统计数据的需求越来越大

2.22. 人口普查数据和调查数据的用户在统计方面共同使用地理数据自动化技术有很多好处。地理信息系统可以把不同课题领域的信息联结在一切。该信息系统提供的数据集成功能已经使统计信息的应用变得越加广泛，因此，统计机构在更大的压力下，不得不为小的地理单位提供高质量的空间基准信息。如果计划周密——也就是说，如果国家统计局办公室能够收集小单位的信息，然后进行适当加总——那么应该能够满足许多新的数据客户的需要。对这类数据的应用实例包括：

- (a) **应急规划和人道主义反应。**火灾、地震、火山爆发或海啸如果发生在人口密度很大的地区，人员疏散十分困难，确定这些地区可以让各个机构对可能发生的自然灾害做好准备。在一场大的自然灾害后，最先提出的问题包括：哪些村庄受到了影响？其人口规模如何？有多少人死亡、受伤或无家可归？基础设施尤其是公路和桥梁、卫生中心、学校、供水系统和政府大楼，等等，状况如何？如果数字化的人口分布和住房特征地图能够与受灾地区的海拔和斜度数据、交通网络以及其他地理信息相重叠，就可以得出受灾人口数、他们在医疗、食品和住处方面的需要，特别是他们的位置的可靠估计。关于居住地的标准的“P-编码”减轻了需要援助的受灾地区的定位困难；
- (b) **洪水平面模型。**大规模水灾正在世界各河流域造成越来越大的危险。利用数字化平面图和水文数据，结合小地域人口普查统计结果，计划制定者可以做出详尽的评估，降低易发生水灾地区人口和应急管理规划的风险。一些国家的保险公司使用这些工具对户主的风险水平进行评估，从而更公平地决定保险费数额；
- (c) **社会和教育服务规划。**地方和地区政府的一个主要任务是，确保全国所有地区能够平等地享受政府的服务，例如卫生和教育方面的服务。通过了解有关年龄和社会特征的小地域人口普查数据，规划者可以对各种不同服务的需要进行预测。结合有关运输基础设施的地理信息系统数据，即可更好地给已有服务中心分配资源，做出有关新设施的合理布局；

- (d) **贫困分析**。在人口普查中没有收集收入和消费数据的国家，家庭特征是不同群体福利情况的重要指数。结合基础设施和农业生态环境的空间基准信息，小地域的人口普查数据可以用于评估贫困的发生率和贫困群体的分布情况。通过把资源注入到最迫切需要的地区，可避免将补贴误用到非贫困群体，因此，这些信息可以增强减贫计划的针对性；
- (e) **公用事业规划**。水、煤气、电力和电信等私营和公营部门不仅可以利用地理信息系统来管理其有形的基础设施，而且可利用人口统计数据的空间分析，评估目前和将来对服务的需求。数字化人口普查数据以及数字化地形模型一直是全世界实地性服务设计的一个关键组成部分；
- (f) **劳动力分析**。无论是对于一个正在寻找合适的地点建厂的私人公司，还是一个正在试图使劳动力供求平衡的政府机构，小地域的人口普查数据在有关雇用的分析中都是一个重要因素。在上班交通问题的分析中，对工作地点和雇员住宅区进行比较，这对于规划交通运输至关重要；
- (g) **市场分析**。公司利用小地域人口普查数据规划建立新商店和仓库的位置，管理顾客服务信息，确定广告对象，从而产生了一个完整的地理信息系统分支(尽管名字不尽相同，或称商业地理学，或称地理人口统计学)。事实上，对于这些不同类型分析的强烈需求一直是开发廉价易用的桌面制图软件的主要动力；
- (h) **选区划分**。在代议制民主政体中，议会的代表性建立在每张选票具有同等分量的原则之上。为确保贯彻这一原则，小地域的人口数字可用于设计面积大体相同的选区。例如，在美国，《宪法》规定每十年进行一次人口普查的主要依据就在于此。地理信息系统和人口普查数据被用于划分选区；
- (i) **疫情分析**。利用小地域人口普查数据，配合健康情况数据和生物物理数据，卫生部门的负责人可以估计出可能受某种疾病感染的人数，或者估计出可能成为某种病菌携带者的人数。例如，掌握了全国疟疾和血吸虫病的潜在感染者人数后，决策者就可估计出采取根除措施所需的资源。确认这些危险人群的聚集地有助于开展各项工作，如优先安排和采取干预措施等；
- (j) **农业**。有关农业生态情况的地理信息、生产数据以及小地域食品需求量数据有助于粮食安全问题的分析。许多生态系统脆弱的国家都建立了饥荒预警系统，以避免出现重大的粮食危机。

## E. 投资于地理空间技术：成本和收益

2. 23. 本节探讨为人口普查活动采纳地理空间技术所涉及的成本和潜在收益。有各种选择，从完全集成的内部地理空间基础结构，到使用，比如说，桌面制图软件来仅仅展示结果和发布结果。换言之，对于将地理空间技术引入人口普查工作来说，没有“一刀切”的解决办法。不过，先进的创新技术和较低的进入成本使得这方面的转化无疑是可以做到的，因为全球定位系统技术变得便宜和复杂，并且可以在廉价的台式电脑上使用高端的地理信息系统。适合于任务是任何成本效益分析最重要的原则。

2. 24. 正如我们强调过的，采纳任何新技术都是对现有组织的挑战，只有在仔细考虑成本和收益之后才能采纳。成本的组成部分包括系统设计、软件采购、原型开发、人力资源规划、培训、合同支持、地理数据库设计、过渡成本、收据获取、数据收集、质量控制、系统维护和产品开发。向地理空间数据转换要发生与数字化有关的费用，即将模拟的点查区地图转换为地理空间格式。

2. 25. 为人口普查工作之目的向地理空间系统迁移的好处分为两类：效率收益和效力收益。效率指的是每个输入单位可以获得的输出数量。就人口普查而言，效率意味着用较少的钱做更多的事。效率方面收益包括成本的节省或生产率的提高、时间的节省、地理空间数据产品可靠性和权威性提高、更好的服务、更加准确、一致性提高以及创收。通过适当的规划，人口普查组织本身就可以实现这些收益。

2. 26. 效力收益指的是得益于信息的改进的政策或方案方面的影响，包括由人口普查统计数据的用户实现的社会收益。效力收益包括在一个对地方和地区研究来说更合适的范围进行更好的分析，在更加可靠的信息基础上做出决策，与其他政府机构或非政府组织进行更广泛的数据共享，改善针对广大公众的外联活动。一些社会收益，例如人道主义反应和健康基础设施规划，从上节中提到的例子就可以推断出来。事实上，有效使用地理空间技术可能意味着救助生命并全面增进人类福利。

2. 27. 由于种种原因，很难用量化手段评估使用地理信息系统的成本效益。例如，有许多好处可能不会使为地理信息系统投资的机构受益，而得到好处的却是局外人，例如那些得以使用更高精度和更低成本的产品的人，或者那些可以使用过去根本用不到的产品的人。这也更清楚地说明了“价廉”和“成本效益”之间的差别。绘制普查地图最便宜的选择可能仍是传统的手工操作方法，尤其是在劳动力成本低的国家。然而从社会的角度看，在数字化进程的初始阶段投入较多的资金可能更具成本效益，因为无论从人口普查和统计机构的内部和外部来看，数字化产品将会实现很大的长期效益，发起一场真正的全国性主动行动。另外，没有人能保证人工成本将保持在低水平。

## 方框II.2

## 技术和成本障碍

## 1. 柬埔寨

作为2003年健康调查的一部分，柬埔寨进行了点查区制图和村庄边界制图。人口普查地理信息系统活动是国家统计局人口统计司工作的一个组成部分。与土地管理部和内政部一起对人口普查制图活动进行了协调。30名工作人员被指派进行点查区划界工作，每个点查区不超过120户。还使用了乡村问卷来制成点查区数据库。点查区数据库的功能包括标注、登记、编码、编辑、进行数据输入、核实和清除边界数据。使用了条形码来方便点查区地图的追踪工作。对于2008年的人口普查，国家统计局计划使用以发展信息数据系统(DevInfo)为基础的CamInfo系统，在线发布人口普查结果。柬埔寨估计制图和点查区划界的成本，再加上数据处理的成本，约占人口普查总成本的15%。从资源的角度看，特别费钱的是实地使用的车辆和计算机(以及计算机外围设备)，尽管这些常常是由捐赠者提供的。

资料来源：柬埔寨的介绍，曼谷讲习班，2007年。

## 2. 莱索托

莱索托在2006年进行了它最近一次的人口和住房普查。利用遥感和地理信息系统方面的最新技术进行了点查前的人口普查制图工作。其中包括全覆盖的人造卫星定位及跟踪2.5米自然色卫星图像、全球定位系统和地理信息系统。筹备办公室的工作由5名工作人员担任。工作包括通过图像解读(以及当地知识)进行初步的点查区划界，以及为实地进行的点查区划分工作绘制卫星图像地图。实地划分点查区的工作由8个小组担任，每个小组4名实地工作人员。每个小组都配备了一部车和一个全球定位系统。必要时对点查区边界进行了核实和修订。还收集了其他属性信息，例如地方名称和地方负责人的姓名。实地工作期间收集的信息以数字格式保存，最后利用地理信息系统制作并打印了点查区和管理区地图。划分了总共4 500个点查区。在9个月的期间内完成了点查前人口普查制图工作。点查前人口普查制图工作约占人口普查总预算的5%。大约40%的点查前人口普查制图预算花在了卫星图像上。尽管这是一个很大的数目，但结果却是节省了超过60%的实地工作，因为不再需要为每个点查区绘制单个的草图了。最后，这一现代方式得到了更加优质的产品，并节省了大约20%的预算(与针对草图方式的最初预算相比)。

资料来源：Geospace Inc.的介绍，卢萨卡讲习班，2007年。

2. 28. 对地理信息系统的投资很多集中在前期进行。也就是说，主要的成本需要在一个项目的开始阶段投入，而实际收益可能在项目的长期运作中才得以实现。将传统的地图绘制方法与数字化绘图进行对比可以发现，一个明显的区别是在传统绘图的情况下，每一次人口普查的地图都是手工绘制的。这包含了许多不必要的重复劳动。过去很常见的情况是，人口普查地图绘制纯粹是以项目为基础的。在人口普查之前几年，一个小组就被召集起来，以手工方式快速制作人口普查草图，这些草图只用于点查目的。几年后，为下一次人口普查又开始了这一过程。

2. 29. 为地理空间技术的采纳制定一个长期战略意味着把人口普查地图绘制和地理数据库的开发看作是一个持续的过程，需要长期的核心工作人员对数据库进行定期维护，并且这些工作人员要接受经常培训。在模拟的情况下，成本一般高于收益，因为硬拷贝地图只用于人口普查的目的。在数字化情况



下，初期的巨大投资很可能导致维护和更新成本降低，并获得长期的可持续收益。长期收益可能相当高，因为这一过程的结果是一个数字化数据库，国家统计局办公室和其他政府机构都可以多次使用该数据库。

## F. 在国家统计办公室实施地理空间项目的关键成功要素

2.30. 对于普查工作中特定的地理空间应用，除了可以量化的显而易见的成本之外，尚有许多难以逾越的障碍，这些障碍可能导致项目失败或者无法完全实现所有预定目标。这些问题通常是由于缺乏规划、选择了不合适的软硬件以及组织方面出现差错导致的。通过对地理信息系统实际项目进行调查可以发现，地理信息系统项目取得成功有一些共同特点。以下按优先次序列举了一些关键成功要素，该清单特别适用于国家统计局办公室的工作：

- (1) 详尽的策略、运作和管理规划要建立在对本成本和努力的切实评估之上，要为地理空间业务制定明确的总体目标和具体目标。要有既定的质量控制/质量保障程序。
- (2) 决定在地理空间技术方面投资，目的是满足需要和解决问题，而不是追求技术进步。不要把技术作为一个独立的补充部分，而要把其作为全面的管理策略的一个组成部分。
- (3) 组织内要有一个促进地理空间系统开发的关键人物，要得到高层管理人员的支持。
- (4) 在人力资源方面，能够为相关雇员和管理人员提供培训和支持。能够满足人员配备需要，包括能够留住熟练的工作人员和酌情雇佣承包商。与政府部门内外的销售者、顾问、合作伙伴和客户签订明确的书面合同。
- (5) 完成用户需求评估并以推理的方式确定输出产品，包括明确的执行进度表。拟定长期供资计划，包括费用回收和数据定价策略，包括准确地估计维护费用和相关费用。规定频繁的转折点和输出产品的交付，以鼓励遵守预定的时间表。
- (6) 与其他有关方面签订合作协议，包括空间数据基础结构协议。通过用户网络上详细的文件记录和元数据促进数据共享。
- (7) 系统和产品的可用性。
- (8) 遵守地理标准。
- (9) 使用数据集成方法。
- (10) 使用清楚的协议，例如用全球定位系统来收集和处理数据。

## G. 利用地理空间工具规划人口普查工作

2.31. 本部分对人口普查绘图项目预备阶段的组织工作进行讨论，并探讨了这些工作包含的关键性设计问题，它们对数据库性质以及所支持的应用范围有决定性影响。能否对实际数据进行成功转换取决于合理设计的制度环境和计划周密的行动策略。此处的计划步骤分为机构问题，例如地理支持的组织结构(包括人员配备和与其他机构合作)，人口普查地理学的明确定义，以及地理空间数据库的设计。如图II.1所示，利用组织上核准的方法，这些阶段在一定程度上可同步进行，其中的许多方案取决于数据输入策略。

## H. 需求评估和确定地理选择

### 1. 用户需求评估

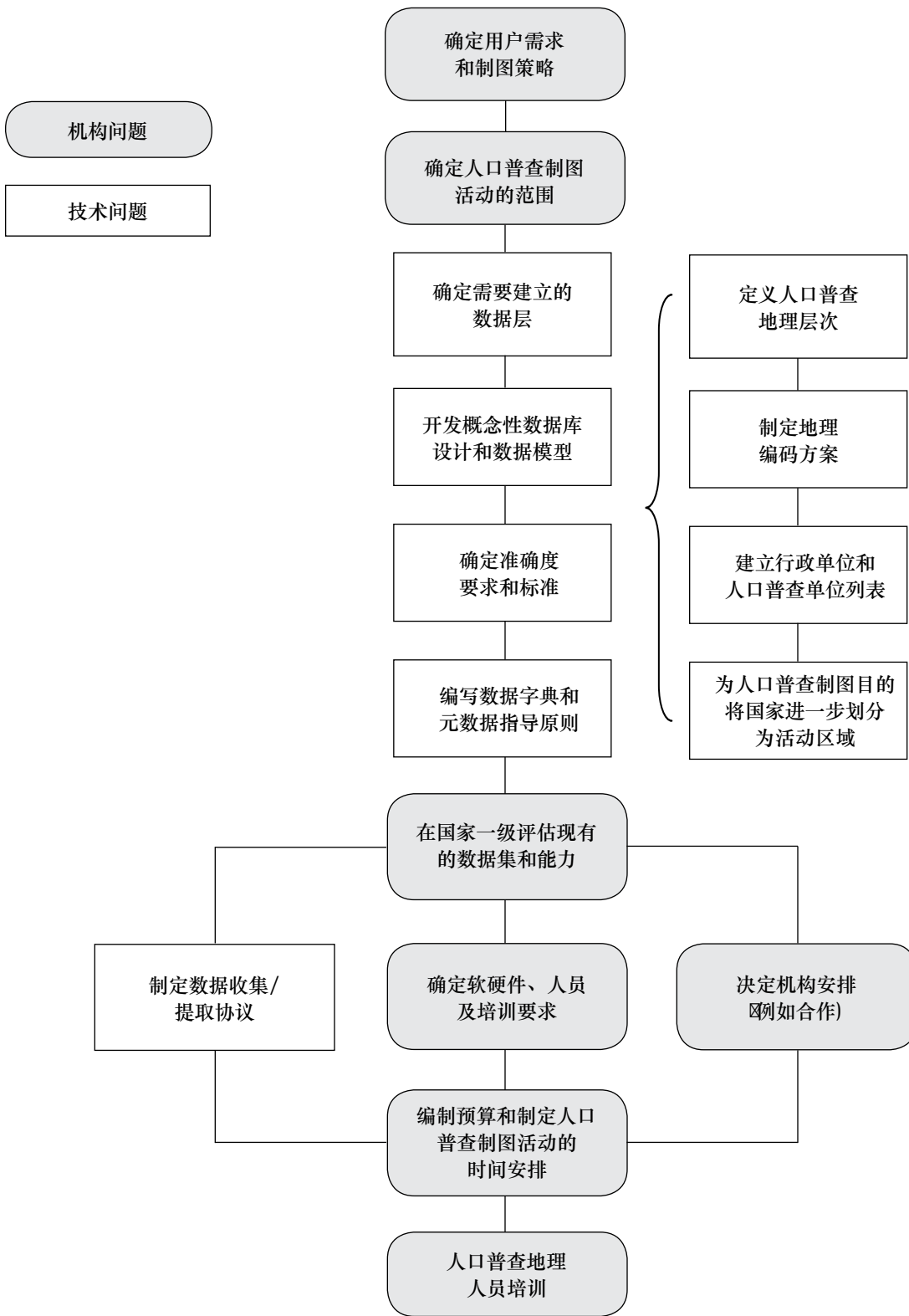
2.32. 人口普查制图项目的第一步是进行详尽的需求评估，然后是地理选择的可行性调查。人口普查制图机构随后必须对用户期望值和切实可行的资源进行协调，从最终产品和服务到需求做反向工作。

2.33. 人口普查规划过程要想取得成功，需要与人口普查信息的主要使用者进行广泛协商。人口普查规划过程应包括就地理内容，即包括行政结构或概要层次在内的地理结构，进行协商，还要就支持人口普查数据分析的地理基础产品进行协商。这一规划部分应包括在针对人口普查的全面协商方案中。由于空间基准人口普查数据的需求增加，有关地理产品的协商将成为规划过程中的突出任务。因此，在人口普查规划过程中，应该把统计地图的使用单位包括在提供建议的咨询小组之中。在规划阶段，人口普查办公室必须与三个主要群体进行协商：

- (a) **参与人口普查活动的人员和机构。**为了获得有关资源和潜在困难的全部信息，国家统计局办公室必须进行广泛调查，掌握有关国家人力资源、可用设备、现有的数字化和模拟式地图产品，以及其他公共和私营实体正在进行或计划进行的相关活动等情况。避免重复他人已经完成的工作是降低人口普查地理活动的成本和及时把人口普查地图产品交付使用的关键；
- (b) **人口普查地理数据产品的使用者。**这些用户主要来自其他政府部门、学术研究团体和私营部门；
- (c) **普通公众。**有了计算机使用和因特网信息展示的各种选择，个人用户也是一个重要的用户群体。例如，居民可能希望获得有关自己小区和准备迁入的小区的统计数据。随着当前技术的快速发展，人口普查办公室必须认真计划，以便对数据需求进行预测。

图 II.1

人口普查地理工作规划中的各个阶段



## 2. 确定产品

2.34. 用户的需要决定了必须在人口普查绘图周期内完成的产品范围。国家统计局制作的产品将在第六章详细讨论。这些产品始终应包括适当的文件记录，包括编码和元数据，以便这些产品对用户最有用。实例包括：

- 一套数字化点查区地图或衍生的发布单位。这些地图旨在实现所有产品的生产，交给各政府部门和公众使用。
- 为所有统计报告单位制作数字格式的地理边界文件，为此，需要将人口普查指数列入表格。
- 列出所有统计和行政报告单位，包括城镇和村庄，其不同的名称以及地理坐标。
- 地理同义文件。这些文件将说明目前的报告单位与原来人口普查使用过的单位的关系，或者一组报告单位与另一组报告单位的关系。
- 包括特征数据的向量层，例如地标、公路、学校、医院和诊所，在对人口数据进行空间分析时可以利用这些数据。
- 所有主要城市地区的街道索引。
- 矩心文件。这些文件给每一报告单位提供一个具有代表性的地理基准。
- 地名字典。地名字典为所有人口聚居地和国家其他重要地理特征提供地理坐标。

2.35. 用户需要是设计人口普查地理数据库的最重要的决定因素。当然，在预算阶段必须根据可用资源对这些需要加以考虑。其他因素也决定了如何选择地理策略，包括：

- 现有的财力和人力资源。
- 现有的数字化和模拟式地理产品。
- 制图机构和统计机构或国内其他相关机构之间的协作程度。
- 统计单位与协作机构的技术能力。
- 技术应用与低技术劳力使用增加之间的平衡。前者可能需要从外国引进，从而增加对外国技术的依赖；后者则可能推进地方经济的发展。
- 国家大小，包括人口和地区范围，以及受地形和水体影响的可达性。
- 可用来规划和开展人口普查制图工作的期限。

## 3. 地理数据的选择

2.36. 各国在自己的现有信息、预算、技术能力和可用时间框架的基础上开始自己的地理工作。因此，有很多方法能够达到为人口普查数据收集和发

布的目的建立全数字化地理数据库的目的。以下列出部分选择，按照复杂性程度由低到高排列，这些方法取决于预算和时间：

- 根据现有的草图制作数字化地图。
- 以地理为基准的点查区层或空间文件，包括适当的编码和元数据，可与其他数字化地理数据库适当集成。
- 使用内含地理基准层的地图。这种地图可以显示诸如道路、河流、地标、点特征和其他特征。这种地图可以通过扫描地图获得的简单图像获得，或设计成单位自己的地理数据库。
- 使用数字化邮政地址登记簿，在这种登记簿里，地址通过自动或半自动手段与数字化道路数据库相对应。
- 利用精确定位的居住单位数字化数据库，该数据库在地理定位系统的支持下建立。

2. 37. 以上列出的内容是不完整的，旨在达到说明目的。所有这些问题都将在本《手册》后面的章节进行详细讨论。一个国家的最佳人口普查地理策略应该考虑到本国需要和资源，量体裁衣。本《手册》将针对可提供的技术和后勤选择进行讨论。人口普查办公室必须从这些选择中，挑选出最适合本国需要的技术和程序。

#### 4. 人力资源和能力建设

2. 38. 有推动力和训练有素的人是一个人口普查地理项目成功与否的关键因素之一。无论地图是由手工绘制还是计算机绘制，人口普查绘图项目的目标是相同的。但是，使用计算机需要人口普查地理人员掌握一系列新技术。因为产品虽然相似，却要用不同的技术完成。另外，数字化地理数据库还可以用于许多其他目的。因此，人口普查办公室很可能需要满足人们对产品和服务的额外要求，这些产品和服务以前是无法提供的。因此，每个人口普查地理人员应该掌握一定程度的计算机使用技能。

2. 39. 在传统的手工人口普查制图方法中，有很多专业知识可用于数字化绘图项目。数字化制图方法不是完全取代已有的技能，而是增加计算机使用方面的专业技能，包括数据库、电子数据表、文件管理和基本网络操作。因此，在工作人员身上，只有较少的制图和地理方面的专门知识过时了，但对其工作技能的要求提高了。例如，接受过传统训练的绘图人员将不再需要使用一些手工绘图的技能，如用钢笔或铅笔书写、画暗线和画草图。经过计算机操作培训后，这些人员将能够利用他们在地图设计和绘图中的优势，通过地理信息系统或桌面制图软件绘制出设计完美的点查区地图或是专题地图。对某个学科的专业人员进行计算机技能培训比起对计算机专业人员进行应用领域的培训来相对要容易一些。

2.40. 下面将详细说明数字化人口普查绘图项目要求人员完成的工作。在人口普查项目的不同阶段，人口普查办公室的同一班工作人员可能需要承担这些任务中的若干项。

2.41. **规划。**在项目的初期阶段，应该组建一个小组来制定数字化人口普查制图的总体策略。这项工作要求小组成员接受地理、地理空间以及计算机应用方面的培训，并且具有人口普查制图经验。除了人口普查办公室的人员外，这个小组还应包括国家制图机构、其他有关政府部门以及数据使用者的代表，或者外聘顾问。规划过程还应有技术顾问的参与，他们来自已经转而使用数字化人口普查制图技术的国家的国家统计局办公室或者国际组织，因为他们能够对规划过程提出有用的建议并有助于成本选择。

2.42. **项目领导。**领导规划过程的人就是人口普查制图的项目领导人，他还应监督数字化人口普查制图策略的执行。他应该在地理、计算机科学或者相关领域内具有一定的造诣，并且受过地理信息科学和数字化制图方面的培训，应该具有人口普查制图经验，最好是从国家以前进行的点查活动中取得的经验。管理经验或管理培训对于指导预算编制、人员管理和日程安排都十分必要。交际能力有助于加强与人口普查项目的其他部门及合作单位的协调。项目领导者还必须紧跟地理空间技术的发展趋势，并且要在情况发生变化或者出现更好的解决方案时，随时对人口普查制图策略进行修改。重要的是一旦过程开始不要制定战略，并且对“任务爬行”要有意识。改变控制机制为应对规划遇到的意想不到的挑战提供了一个不可或缺的手段。

2.43. **地理空间数据的转换。**数据转换专业人员是负责实际完成地图信息向数字化数据库格式的转换工作的人。他们接受过相关技术的培训，例如，通过关系型数据库管理系统进行扫描、数字化处理和编辑地理空间数据库以及归纳数据库的发展等。数据转换专业人员必须确定数字化数据库开发和指导技术人员的最佳方法。了解新数据源和获得新材料来帮助创建地理数据库的所涉技术同样非常重要。

2.44. **地图扫描和数字化。**虽然扫描有了新的转变，并且在许多国家统计局办公室的数据输入策略中占据了突出的位置，但数字化仍是一种可靠的选择。没有在地理或类似领域受过专业培训的人能够相对迅速地掌握这一技术知识。然而，数字化处理是一项高度重复性的工作，要求高度专心，注意细节，要对数字化地理数据库的结构有清楚的了解，通常可以被分配给文员去做。优秀的数字化处理人员还应该接受有关质量控制/质量保障方法的培训。这一培训要求同样适用于平视数字化处理。

2.45. **绘图设计。**绘图人员将负责设计所有的地图产品，包括点查区地图、监管地图以及反映普查结果的专题地图。他们必须懂得地图设计和地图表现方法，必须接受地理信息系统和数字化制图方面的培训。受过训练的绘图人员具备所需要的大部分技能，但是在计算机操作方面必须接受足够的培训。

2.46. **现场工作。**随着数字化地图生产技术的使用，对人口普查绘图现场工作的要求也发生了改变。因为全球定位系统已经成为现场数据采集的主要工具，现场工作人员不但必须学会在工作中使用这些系统，可能还要学会通过笔记本电脑在现场下载和显示这些数据。虽然无需掌握地理学或调查方面的专业背景知识，现场工作人员就必须接受正确使用新型工具的培训。

2.47. **系统管理。**前面已经提到，国家统计局办公室需要投资于专门的信息技术中心，该中心能够满足地理信息系统以及其他信息和通信技术/数据处理需要。能否按时完成人口普查地理数据库项目取决于能否使计算机设备正常工作。能否及时完成人口普查地理数据库项目取决于计算机设备的顺利运行。一名系统管理员负责维护计算机硬件和软件系统，其目标是尽量减少停机时间，支持人口普查绘图人员以及确保数据的安全(例如，数据备份)。系统管理人员不直接参与人口普查制图工作，尽管如此，他们却是地理部门的重要成员，因为，几乎各方面的工作都要依赖计算机系统的正常运转。在某些情况下，人口普查办公室地理部门的计算机系统管理工作可由机构内的通用计算机支持人员承担。

2.48. **特殊要求。**根据采取的人口普查制图策略，人口普查制图组织也许还需要更多的专家参与。例如，如果人口普查地图的升级工作非常需要利用遥感产品，那么人员中就应该包括训练有素的数字化图像分析专家。可能需要的其他专家包括：大容量地图扫描系统操作员，或熟悉数据库管理软件系统以及计算机编程的人员。还有一些人可能具有图形方面的能力，web程序部署或者客户关系方面的技能。所有这些技能在开发数据库和优化软件系统方面都是有用的。

2.49. **各种层次的培训。**在长期或临时参加人口普查制图工作方面，很多国家可能缺乏训练有素的地理空间专家。因此，人口普查办公室必须对培训方案进行评估，确保已有人员和新手具备成功完成项目所需的适当知识。一般地说，经过培训后，那些具有一定计算机知识和受过传统地理技能培训的人不难适应数字化技术的要求。为不同目的需要进行的培训包括：

- (a) 为人口普查办公室的全体人员，包括来自其他部门和管理层的人员，举办短期讲习班，提高他们对人口普查数字化制图方案的了解。这有助于数字化制图项目更好地融入人口普查的全过程。广泛发布信息的另一好处是，能够有效地利用其他人口普查办公室制作的人口普查地图产品。此类讲习班可以由人口普查项目领导或人口普查制图部门的专家主办；
- (b) 对诸如扫描、数字化处理或编辑这样的重复性工作的培训可以举办短期不脱产培训班，随后进行岗位培训。新手完成的产品应该接受严格审查，以便确定人员是否需要进一步指导或培训，是否有必要调换到其他岗位；
- (c) 参与人口普查制图工作的核心地理专业人员应该接受更多的有关地理信息系统软件和数字化制图技术方面的培训。因为培训费

用很高，所以只能将长期工作人员送到国内外大学、开发商和其他组织举办的学习班学习。受过此类培训的人应该在信息传播和培训其他员工方面发挥重要作用。大量的人员培训可以通过分级“培训教师”的方法开展，这种方法有其适合分散式人口普查绘图工作；

- (d) 专业技能的应用，例如数字化图像处理或先进的计算机数据库的应用，通常需要一定的专业技术水平或同等的实际经验。如果雇用不到合适的人员，人口普查办公室应该早在制图项目实际开展之前就考虑派一个人到大学接受培训。现在，世界各地有若干所大学和培训中心专门开办了一年至两年的，有关地理信息系统和遥感以及相关技术的学位课程(提供地理空间信息和技术领域的培训的机构名单见附件七)。

## 1. 机构合作：国家空间数据基础结构： 确保与其他政府部门的兼容

2. 50. 在很多国家，几个政府机构都创建了数字化地理数据库。国家制图机构越来越多地在整个地图绘制过程中使用全数字化技术。但是，其他政府部门，包括运输、医疗环境和水资源管理单位，也适用地理空间技术处理他们收集的信息或者用于分析和规划。另外，私营部门的公司，如公用事业、电信和采矿部门的公司，也认识到数字化地理模式为他们的信息管理带来的各种好处。

2. 51. 为地理信息建立基础结构被认为对会员国来说与建立公路、电信网络和提供其他基本服务一样重要。如今得到广泛承认的是发展国家数据基础结构将更好地促进政府机构、私营部门、大学以及民间社会对空间数据的利用和获取。事实上，空间数据基础结构使得国家统计局办公室可以在国内获得专门技术以及现有数字资源，例如用于地理信息系统软件应用程序的地理空间格式的基本数据。在该框架下，国家统计局办公室被看作是数据网络中的节点以及覆盖全国的内容提供者。

2. 52. 如今在全世界100多个国家都有国家空间数据基础结构。一些可以帮助国家统计局办公室建立其空间数据基础结构的可能的信息和专门技术来源包括国家制图当局、环境和规划部以及军队。国家空间数据基础结构的基本组织结构包括一个负责的部、一个主管机构、数据生产者和用户的论坛或网络、一个筹划指导委员会以及多个技术工作组。

2. 53. 与现有的国家空间数据基础结构建立联系需要确定主管机构的联系人。主要的联络机构一般是国家制图机构。可以共享的数据例子包括扫描的地形图或比例为1:100 000或更小的数字化向量图的数据库，数字化向量图可以作为输入物用于建立点查区边界、海拔数据、水文图和运输网络——所有这些都可以帮助进行区域划界，以便进行人口普查-民意调查。



### 方框II.3

#### 数据共享合作的三个例子

##### 1. 人口基金和斐济的硬件共享

用于人口普查地理的新技术的发展为提高人口普查工作的效率提供了强大的工具。在斐济，来自斐济统计局的一个小组进行了实际点查工作，他们利用全球定位系统装置对有人住的每个住宅单位进行了地理标注。这是如何给人口普查增加价值的例子之一，因为这一程序使得未来的绘图应用更为准确，并将为未来的人口普查和统计活动提供便利。为全球定位系统演练提供资金得到了联合国人口基金(人口基金)的帮助，人口基金提议将全球定位系统装置的使用扩大到想要进行相同或类似演练的其他太平洋岛屿国家。这一共享协定对于能力建设和在人口普查期间利用宝贵的信息来说非常有价值(如果想了解更多信息，请联系Scott Pontifex, 太平洋共同体秘书处)。

##### 2. 太平洋岛屿应用地球科学委员会秘书处和太平洋岛屿国家的数据共享和处理协定

与几年前相比，获得卫星图像的费用低了很多，这些卫星图像使得能够将地理信息系统应用软件用于人口普查地理活动，如点查区划界。不过，对于大多数国家统计办公室来说，这些费用仍然难以承受。国际和国家机构可以大量贡献其现有的数据集，为开展人口普查活动提供帮助。在该背景下，太平洋岛屿应用地球科学委员会秘书处获取了卫星图像，它可以将这些图像免费提供给所有太平洋岛屿。该委员会还可以提供现有地图的扫描和地理登记。这对于在一个当代平台上发展人口普查地理来说是一个很不错的起点。各国的统计办公室正在考虑利用该委员会提供的这些选择来评估利用地理信息系统进行人口和住房普查(如果想了解更多信息，请联系Scott Pontifex, 太平洋共同体秘书处)。

##### 3. 美国的数据共享经验

美国人口普查局与部落、州、县和地方官员以及协调机构，例如区规划委员会一起开展工作，以实施为共享地理信息而设计的几个方案。地方合作伙伴提供的知识使美国人口普查局能够满足统计和空间数据方面的需要，这是其使命的一部分：作为统计数据的来源，让人们更好地了解美国。

美国十年一次的人口普查和一年一次的美国社会调查主要通过向全国住户邮递调查问卷的方式进行。与美国邮政局的密切合作使得可以对地理司的地址主文件进行定期的地址更新，该文件在邮递调查问卷时使用。根据保密协议，美国人口普查局与部落、州、县和地方政府官员共享其地址清单，以确保邮递的准确性，从而保证计数的准确性。

正确的边界位置对于收集、汇总和发布统计数据十分重要。美国人口普查局每年进行一次边界和合并情况调查，在调查中，地方政府对边界进行审查，并报告任何变动情况。在筹备十年一次的人口普查期间，区域规划机构和政府有机会审查统计区并提出更改建议以更好地满足地方数据的需要。

地理信息系统在美国地方政府中的发展为像美国人口普查局这样的联邦机构通过合作伙伴关系获取准确度高的现有地理空间数据提供了机会，这些数据可以帮助该机构维持其全国性的拓扑集成的地理编码与参照系统(MAF/TIGER)。MAF/TIGER是统计机构所有地理活动的源泉，例如提供支持、空间数据的使用、地理编码和制图。

在建立一个国家空间数据基础结构的工作中与其他联邦机构的协调和伙伴关系是通过美国人口普查局与联邦地理数据委员会之类的团体的积极参与实现的。与国家测地局之类的机构的协议有助于该机构获取更准确的关于居住单位的位置的信息。最后，与国家各州地理信息理事会和国家各县联合会之类的全国性组织的密切合作有助于与地方政府建立必要的伙伴关系。

合作伙伴关系方面的当前趋势对于维护进行人口普查和调查所需的地理空间框架至关重要。在美国，建立和维护与过去几十年依靠联邦的集中努力的地址数据和空间数据同样数量和质量的地理空间基础结构的更好的数据提供了新的机会(如果了解更多信息，请联系 Tim Trainor: timothy.f.trainor@census.gov)。

\* \* \*

关键是在需要获得、使用和开发地理信息系统应用软件时，确保国家一级的合作。建立地理信息系统协调中心和用户群体是一种最佳方式，可以确保协同作用，避免重复劳动，浪费宝贵的人力资源。国家和国际机构提供持续的和安排得当的技术援助的重要性怎么强调都不过分。对许多国家来说，这是为开展复杂的工作，例如人口普查进行的能力建设的一个重要组成部分。人口普查成功与否，常常取决于人口普查之前的公共宣传和提高意识活动是否成功。用于开展这样一项活动的资源常常来自外部捐赠者和国际机构，必须充分利用这些资源。

2. 54. 政府机构内外的众多用户需要使用基本地理数据库。其中许多用户需要同时访问几个数据库，或者利用一个标准地理数据层作为其收集自己的空间数据的模板。许多地图绘制和数据收集活动在这些标准数据层的基础上展开，这些标准数据层被称为“框架数据”。构成国家空间数据框架要素的核心数据层包括(见非洲经济委员会(2007年)):

- (a) **大地测量控制。**一种精确确定地理控制点的系统，起到为国家所有地图绘制活动提供基准的作用，有时被称为基准；
- (b) **基础地理。**包括图像(航摄照片或高分辨率卫星图像)、地势图(地形)以及水文图(地表水特征；可以是自然的，例如河流和湖泊，也可以是人为的，例如运河)；
- (c) **行政和空间组织。**包括在地面划分边界的国以下各级政府单位，例如省和区，地理名称和土地管理单位/地区；
- (d) **基础设施。**包括公路、内陆水道、铁路和用于运输人员或货物的任何基础设施、公用事业和服务；
- (e) **自然环境。**包括土壤类型、植被带、人口定居点的地理名称。

2. 55. 与人口普查办公室打交道的大都是政府行政单位。显然，点查区需要与国家行政区划的边界和人口分布保持一致。但诸如运输和水文图这些数据层对人口普查制图也很重要，因为道路与河流为点查区形成了天然轮廓。

2. 56. 在进行人口普查后，国家统计局办公室应与国家空间数据基础结构的其他参加者共享人口普查的结果和产品。点查区边界，以及人口普查信息，例如基本的人口信息，也许还有住房和社区特征，对于其他政府和私人组织来说都是重要的数据来源。国家统计局办公室还可以提供一个包含点查区中心点(矩心)的数据集。例如，健康部门的分析需要关于高危人群的详细信息。运输部门的规划可以很好地利用关于对公共运输服务的需求的数据。另外，公共和私营公用事业需要了解应在哪里提高电、水或电信服务的供给能力。

2.57. 国家空间数据基础结构包括以基本地理基准为标准的地理数据库。这一概念说明人口普查制图活动包括三个方面：

- (a) 人口普查办公室负责给国家空间数据基础结构提供一系列数据报告单位，这些单位必须与行政结构一致，并能与之沟通社会经济信息及有关信息。为了确保这些人口普查地图能与其他数据资源集成，人口普查制图机构应坚持使用现有的国家地理数据标准；
- (b) 为确保与其他数据兼容和促进人口普查地图的完善，人口普查制图单位应该和其他与地图绘制工作有关的政府部门密切合作。除了确保采用一致的标准和定义外，由于能避免重复性工作，合作还有利于降低成本；
- (c) 国家统计办公室应特别注意与地理基准有关的问题。为了覆盖包含地理范围相同的地区的不同数据集，需要了解不同特征的位置是如何界定的，也就是说要知道坐标系的投影和数据以及细节，以确保不同数据集中地理特征之间的正确的空间关系。

2.58. 一个统计机构对全国性的国家空间数据基础结构工作的贡献还可以包括其他参与形式，比如承诺出席规划会议以及随时了解国内进展。国家统计办公室还需要提供元数据。

#### 方框II.4

#### 国际机构的参与和协调

##### 1. 全球地图倡议

全球地图倡议是1992年在巴西里约热内卢举行的联合国环境与发展会议上提出的。1996年建立了全球测图国际指导委员会，目的是协调参与国的官方绘图组织网络。到了2007年3月，全世界172个国家和地区都在为全球地图提供数据或开发数据。全球地图的目标是以1:1 000 000的比例覆盖地球的整个陆地面积，其空间分辨率为1 000米。大约每隔五年将对数据进行一次更新，以监测随时间发生的变化。全球地图是一个平台，支持光栅格式的四个底图(土地使用、土地覆盖、植被和海拔)以及向量格式的四个数据图(人口中心、灌溉、运输和边界)。关于全球地图的数据通过因特网向所有人免费提供，用于非营利目的(如果想要了解更多信息，见<http://www.globalmap.org>)。

##### 2. 发展信息数据系统

联合国儿童基金会(儿童基金会)发展信息数据系统是一个软件工具，帮助各国监测千年发展目标，并支持通过政策措施、跨部门战略和制定适当的干预措施实现这些目标。发展信息数据系统是一个用于编制和显示数据的通用软件包。除了提供数据库外，发展信息数据系统还提供简单易用的实用工具来查询数据库和制成图形(表、曲线图和地图)，以便包括在报告和介绍中。发展信息数据系统有望成为一种强大的支持手段，促进在国家一级提高对千年发展目标的认识和了解以及基于证据的决策(如果想要了解更多信息，见<http://www.devinfo.org/>)。

### 3. 联合国空间数据基础结构

联合国空间数据基础结构是2005年提议的，目的是促进和实现可持续发展以及改进人道主义和维和行动。对联合国空间数据基础结构的设想是一个综合、分散的地理空间信息框架，通过使得能够以快速和安全的方式访问、检索和发布地理空间数据和信息，为各级决策提供便利。联合国空间数据基础结构使得联合国机构、有共同利益的联合国机构团体和联合国会员国及其区域和主题分组及合作伙伴中为特殊目的开发的地理空间数据基础结构能够协同工作。有了联合国空间数据基础结构就能够访问、检索和发布地理空间数据和服务，避免联合国内部的重复工作(如果想要了解更多信息，见<http://www.ungis.org/unsdi.htm>)。

## J. 标准

2. 59. 为使数据用户之间能够更好地进行数据交换，显然有必要协调地理数据库的开发工作。为此，有些国家成立了国家地理数据委员会，其成员为负责空间数据开发的关键人员。另外，有些跨国组织积极制定地理数据标准，如欧洲地理信息保护组织(EUROGI)、全球空间数据基础结构(GSDI)协会、国际标准化组织地理信息/测绘委员会(ISO-TC/211)和开放地理信息系统联合会(OGC)。

## K. 合作

2. 60. 在创建数字化地理数据库的过程中，鼓励人口普查单位与其他政府机构合作或者与私营部门合作。两种选择在不同的国家都曾取得过成功。如上所述，在政府机构中，首先需要接触的是国家制图部门，尤其是对刚刚起步的国家统计办公室而言。但其他机构(例如地籍、环境和地方政府机构)也许能够提供资源或愿意分担建立高质量人口普查数据库的费用。在私人机构中，软硬件销售商可以在人口普查制图过程中给予技术支持，既可以与人口普查办公室签约，也可以签订成本分担协议书。私人公司的投资将通过出售空间基准数据库得到补偿。其他合作者可能包括数据提供者、学术部门和公路管理机构。但必须注意，与其他机构的合作只是一种愿望，而非硬性规定。既然人口普查制图机构的主要优先任务是为人口普查部门提供地理基础，它必须避免依赖外部的数据供应者。

2. 61. 任何合作或协作必须建立在共同意愿以及内容明确的协议基础上。合作协议或谅解备忘录应该明确规定以下内容：

- (a) **合作形式。**松散的合作关系就足够了，还是要将各项安排严格地确定下来？制定一份更加正规的协议需要花费较多的时间，但可避免以后在数据产品的开发和使用权利及责任方面出现异议。因此，在大多数情况下，人口普查机构与合作单位之间应当制定一份正式的和具有法律效力的谅解备忘录，其中要包括所有与合作关系有关的问题。在与私人数据或服务提供商打交道时，必须订

立这种具有正式契约性质的协议书。要考虑的重要问题包括制定供内部和外部使用的标准的必要性、明确提出的目标的意义以及管理支持的必要性；

- (b) **合作范围。**合作协议可以只包括使用另一个机构的数据，但可能也会涉及一个大规模综合空间数据库从无到有的开发工作。为了利用来自市政当局以及其他较低级别的地理或行政实体的数据，应拟定协议；
- (c) **责任。**由谁负责哪些任务？发挥哪些作用？需要讨论的问题包括数据开发、维护、访问、项目指导以及资源利用；
- (d) **收益。**显然，项目安排必须有益于各个合作方，除非某一个机构完全是购买另一个机构提供的服务。为了公平分配任务和责任，需要明确指出各合作方将如何从项目安排中受益；
- (e) **资源需求。**资源包括人力资源、计算条件、材料以及通信。管理和项目指导所需的资源也必须考虑在内；
- (f) **费用分担。**任何与合作有关的直接和间接费用都应合理分担。账面上不一定能做到一览无遗，因为贡献能以现金、数据、劳力、设备使用或其他形式存在；
- (g) **成本回收。**如果最终产品销售获得收入，应该共同分享。收益的分配应考虑管理和数据销售成本。同时应明确产品的协议使用权和版权；
- (h) **矛盾的解决。**在完成项目的过程中，如果发生分歧，事前需要有矛盾解决方案。

## L. 概要和结论

2. 62. 对地理空间人口普查进行规划需要考虑供资、人员配备和项目管理问题。正如前面所论证的，国家统计局办公室有很大机会利用新的地理空间技术来获得关于人口的信息，包括可以普遍获得个人电脑、便携式设备、全球定位系统和低成本的航摄照片和卫星图像。对许多国家统计局办公室来说，任务是获取新的才能并重组机构以适应新技术。这意味着要认识到在吸引和留住训练有素的工作人员方面面临的挑战，并且许多国家对如何解决这些问题表示关切。

2. 63. 国家统计局办公室需要了解在国家空间数据基础结构方面开展的工作，这既是为了获取宝贵的基础数据，以便用于人口普查规划，也是为了在进行人口普查后分享数据和信息。这里的观点是促使国家统计局办公室从承诺阶段进入操作阶段，并且开始解决实际问题。国家统计局办公室因而将能够获得它们需要的答案，以制定一个对国家来说规模合适的地心数字化计划，该计划将帮助它们实现效率和效力方面的收益。应该铭记的是规划极为重要。机构领导层应明确提出目标并支持战略计划以确保成功。

## 第三章

# 为人口普查建立一个 点查区级别的数据库

### A. 引言

3.1. 正如第二章中指出的，利用地理空间技术来生成更好的数据在某些方面是一个组织问题，涉及目标的制定和利用合适的人力资源技能。围绕一个地理信息核心重组国家统计局意味着包含一个国家的地理和国家统计局使用和生成的各个系列的信息之间的关系。地理与数据库之间的关系是通过编码机制实现的。第一步是将第二章中描述的管理资料与技术内容联系起来，表明地理人口普查数据库是如何成为存储和访问各种形式的人口普查信息的活动的中心的。

3.2. 本《手册》的一个重要目标是为建立一个人口普查点查区级别的数据库制定一个操作计划，反映进行人口普查的方式发生的变化。人口普查不只是一种数据收集活动。对会员国来说，它还是一个在提高信息技术能力的同时促进其发展目标的机会。尽管以前人口普查是一种一次性的活动，通常10年开展一次，为人口普查维持一个准确的地理数据库如今是一个持续的过程。

3.3. 通过转向数字化地理基准活动来支持人口普查工作在未来将会取得丰厚回报，因为公众有更多的机会得到新一代产品。通过采用地理空间技术朝准确进行人口普查的目标迈进是服务于公众的善举。由于重新设计人口普查地图绘制活动需要投入精力和资金，国家统计局可以借助前人的工作，并利用前些年绘制的地图。

3.4. 本章将介绍地理编码的概念，这个概念与地理信息系统行业通常理解的指地址匹配的概念略有不同。联合国定义的“地理编码”更加宽泛。它指的是统计观察与用经纬度或其他位置属性表达的现实世界中的位置之间的联系。简而言之，地理编码是一种确保数据知道它们在哪里的方式。虽然本章提出了这一更为宽泛的地理编码的定义，但它也强调了用于人口普查的传统编码方式(即属性编码)的重要性。

3.5. 地理编码的用意是覆盖一个连续的空间范围，从个别住宅单位到点查区级别，再到较高的行政或国家级别。其成功使用取决于一个国家是否建立了一套行政区域，这些行政区域的版图是已知的，并且以计算机编码的形式予以了数字化展现。编码方案必须反映出地理编码方案的一个重要性质，那就是可量测性和灵活性，即全面覆盖，又有增长的空间。

3.6. 行政级别所依据的观点是在国家领土内有用于划分州或省及地区级别的真实土地范围，或者用于选举或健康监测或邮递目的。这些不同的地理特征都可以储存在一个带行政级别编码和单位数量的数据库中。例如，行政级别二的单位(ADM 2)是省，而行政级别三的单位(ADM 3)是区。理想的状态是，任何地理空间活动都使用了地理信息系统格式的这些单位，用于其各种项目。

3.7. 本章随后将把重点放在作为一个管理问题的标准和协调上。国家标准需要协调。如果统计机构是权威机构(即代码的管理人)，那么国家统计局办公室有必要制定一个透明的和合理可辩证的计划。随后，必须强调区域协调的必要性，尤其是在大国。

3.8. 国家统计局办公室在行政边界划分中的作用随国家而变化。国家统计局办公室可能拥有确定国以下地面单位的法定权力，或者它或许被认为拥有事实上的权力。如果国家统计局办公室不是国以下边界的管理人，那么它将不得与其他政府机构一起开展工作，以利用现有的地理编码计划，有可能要修订该计划以满足特殊的人口普查需要。某些区域协调也可能是必要的。

3.9. 本章涉及的主题包括国家人口普查地理的定义，包括行政级别和标准以及点查区划界过程；包括点查区在内的地理区域的编码，以及确保其与以前的人口普查兼容；用于点查区划界和输入程序的地理数据来源；通过扫描和数字化进行的地理数据转换；建立和保持拓扑结构；执行一个点查区数据库；数据质量问题；以及元数据的开发。

3.10. 同样应审慎地制定计划。此外，必须强调的一点是，随着对小区域数据的需求增长，有必要评估国家的实际需要。

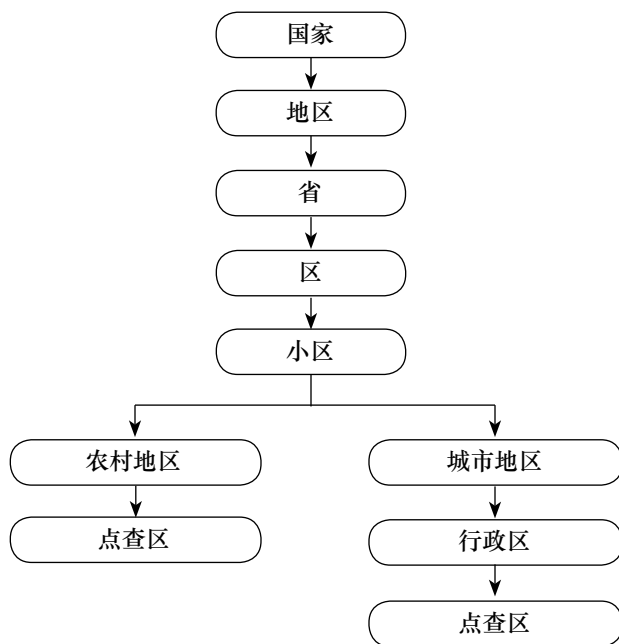
## B. 国家人口普查地理学的定义

### 1. 行政结构

3.11. 人口普查规划最早做出的决定与人口普查数据需要汇报的行政区域有关。行政区域可以是任何特殊的地理单位，但主要是行政单位，也就是说一些政府机关对地域有管辖权。人口普查的准备工作包括制定一份国内所有行政管理机构和统计报告单位的名录，以及明确各种行政管理机构和报告单位之间的关系。每个国家都有其具体的行政结构，即，国家和行政管理单位的下级机构(最低一级除外)被进一步分为更低一级的机构。例如，出于人口普查的目的，一个国家可能将城市地区分为七个等级，将农村分为六个等级(见图 III.1)。

3.12. 在这些等级中，可能只有一些具有实际行政管理功能，例如，省、区一级可以有首府，并拥有管辖那些地区的地方政府办公室。其他单位可能只有统计职责，也就是说，设计它们的目的是为了显示数据，而不是为了管理辖区。图 III.2 用一个仅有四个等级的简单实例说明了行政管理和人口普查单位相互嵌套的关系。然而，在某些实例中，行政管理单位可能无法完全嵌套

图III.1  
一个普通的人口普查地理结构



在一起。当把行政管理单位和其他人口统计汇报单位同时考虑时尤其如此，人口普查办公室可能需要与一个非常复杂的地理区域系统打交道。

3. 13. 并非所有的级别都同等重要。例如，许多国家将领土划分成一些主要区域，这通常是从地理的角度界定的，例如“北/南/西南/东”，或者“山区-平原-沿海”。这些区域通常不具有任何行政管理功能，但仍被用来进行人口普查资料汇报。

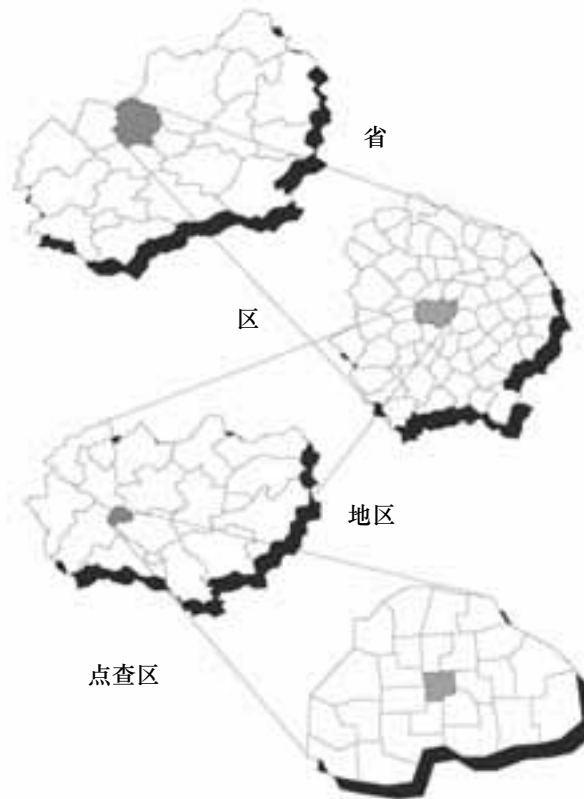
## 2. 行政管理与其他统计汇报或管理单位之间的关系

3. 14. 除了行政单位之外，许多国家还有不少用于不同目的的其他区域，需要为其编制人口普查数据。这些区域有特殊用途，也许是粗略界定的。一些例子如下：

- 健康区。
- 劳动力市场区。
- 选举区。
- 邮政区。
- 文化区或部落区。
- 城市聚集区或大都市区。
- 农业或经济普查单位。
- 土地命名或地籍单位。
- 公用事业区(水或电供应区)。



图III.2  
嵌套式行政结构图示



3. 15. 通过与国内负责机构互动或者通过空间数据基础结构活动可以对这些特殊的地理区域有更多了解。在其空间再现中，一些区域也许无法完全套入该国的行政管理体系。在设计点查区时，人口普查制图机构应当尽量考虑到这些汇报单位，以便利这些区域的人口普查数据制表工作。应在人口普查规划阶段进行用户需求分析时制定出指导原则，根据这些原则决定哪些非行政管理区应该得到最多考虑。一般来说，为了指导点查区的设计，人口普查制图机构应当把所有类型的区域按必须兼容、尽量兼容和不可能兼容三种类型进行划分，并分别加以考虑。

3. 16. 对于国家的某些汇报或管理区来说，数字化边界数据可能已经由各负责单位完成。例如，一些已经开始进行土地改革计划的国家正在利用地理信息系统进行土地命名数据库(地籍资料)的管理，并且许多国家的邮政组织正在使用地理信息系统的邮政编码数据库来方便邮件的投递。那些拥有数字化数据库的单位可以对人口普查地理数据库的开发提供支持。如果能获得很高的兼容性，还会给其他区的统计带来很多额外好处，例如，水的需求量和选举结果可以很容易地与人口统计和社会统计结果结合在一起。

3. 17. 在统计办公室里，其他人口普查操作也需要对数据收集单位进行界定。更为重要的是，许多国家定期对农业和经济状况进行普查。许多分析应用从人口普查信息与农业或经济数据的结合中获益。地理单位之间用于编译这些数据类型的高层次协议大大地增加了它们在公共和非政府领域中的应用。

3. 18. 关于谁管理行政边界以及什么时候能做出改变，统计办公室与其他机构或各级政府之间也应保持协调。在人口普查之前，应冻结边界(在人口普查之前至少六个月建议这样做)，这样的话，边界的改变不会在实际进行区域编码的工作中造成差异。国家统计局办公室还可以选择给特殊版本的行政边界打上“时间戳”，以反映点查时国家内部的划分。

### 3. 点查区地面划界的标准和程序

3. 19. 点查区是用于收集人口普查数据的有效地理单位，在人口普查工作的早期就确定了。不论是使用手工的还是数字化的绘图技术，点查区的划界工作是相似的。点查区的设计应考虑到各种标准。进行了正确划界的点查区应满足以下条件：

- 点查区之间应当相互排斥(互不重叠)并且做到详尽彻底(覆盖整个国家)。
- 地面上应有易于分辨的界限标志。
- 应当与行政结构一致。
- 应当紧凑，无真空或脱节现象。
- 其人口规模应大致相当。
- 足够小和方便，在人口普查期间点查员可以跑遍点查区。
- 足够小和灵活，以便为不同的统计报告单位进行最广泛的制表。
- 满足政府部门和其他数据用户的需要。
- 可以用于其他类型的人口普查和数据收集活动。
- 应当足够大以保证数据的隐私性。

3. 20. 在这些标准中，有些是为了方便人口普查数据的收集，而其他标准则与点查区输出产品的用途有关，即数据收集和制表单位之间的关系。应当切记的是，人口普查的目的是为了给行政管理人员、政策制定者以及人口普查数据的其他使用者提供有用的数据。因此，与点查的便利性相比，在输出最好的产品方面，应首先考虑获得最大的灵活性和适用性。但是，点查区划界必须逻辑上合理，有利于实地操作。

3. 21. 点查区的大小可以通过两种方法确定：以面积或以人口确定。对人口普查地图绘制而言，人口的数量是一条更为重要的标准，但是表面积和交通是否便利也应当加以考虑，以确保点查人员在给定的时间内能够完成点查区的工作。选择人口的数量因国家的不同而不同，一般根据预先测定的结果决定。城市和农村的平均人口规模也不一样，因为城市和乡镇的点查工作要快于

农村。在特殊情况下，可以规定大于或小于平均值的点查区。从最实用的目的出发，点查区的人口数量应当在一、二百至四、五百之间。

3.22. 在划定点查区边界之前，应该估计一下居住在该地区的人口数量及其地理分布情况。除非能从最近的调查、登记系统或者其他信息来源获得资料，否则应该通过清点住户、确定家庭的有关数量和乘以平均住户数量的办法确定人口数量。通过地图绘制的实地工作、与政府官员合作、从以前的人口普查结果推断或者借助于下一章将讨论的航摄照片或卫星图像，可以确定住宅单位的数量。

3.23. 点查区的地面边界必须清晰可辨。所有的点查人员，即便他们没有经过多少地理方面的培训，也应当能够在他们所管辖的地区找到这些边界。因此，为了绘制出明确的轮廓图，各点查区的人口数量会有所不同。可以用于这一目的的自然特征有道路、铁路、小溪和河流、湖泊、篱笆或任何其他可作为边界线的明显特征。边缘比较平缓的特征，如灌木丛、森林或山脊一类的等高线则不很理想。行政管理边界常常不明显。使用地面边界并不清晰的点查区边界难以避免。在这种情况下，需要在点查区的地图上加上确切的文字说明和适当的注释。这方面的例子是支线和延长线。例如，某条点查区分界线沿某条明确规定的支线与某条道路平行。或是某段点查区边界线可能定为从某条可以明显看到的道路延长到另一个明显的特征，如河流或铁路。

3.24. 很多国家会遇到与点查区划界有关的具体问题。例如，某些村庄划归某行政管理单位管辖，却没有明确画出村庄范围的实际边界。而且，有些特殊人口，如暂住人口、盲流或军事人员，都需要参考地理因素来决定。例如，海军通常以他们自己的港口为所在地。在规划难以点查的人口的定位时，应该考虑到这方面的业务费用比起城市地区的居住人口来有时要高出10到20倍。

3.25. 点查区划界标准之一是确定理想的点查区大小，其依据是一名点查员在收集数据的预定期限内能够计算的人口数。点查区划界计划应反映整个人口普查计划，首先是分配给点查工作的天数。一次人口普查预备调查可以确定一名点查员每天能够调查的住宅单位的数目。比如说，如果在城市地区每天可以点查16个住宅单位，而在农村地区每天只能点查10个住宅单位，并且如果点查期限为10天，那么理想的的城市点查区应包括160个住宅单位，而农村点查区应包括100个住宅单位。如果一个住宅单位里包括的人数为5，那么对于一个城市区域来说，理想的人口规模为800人，对于一个农村区域来说则为500人。影响一个点查区大小的其他因素包括行政区边界、点查区边界特征的可见性、集体生活区的存在(例如兵营、旅馆和学校宿舍)以及运输方式和可获得性。

3.26. 人口估计数对于正确划分点查区边界至关重要。可以要求地方官员提供小区域估计数，或者国家统计办公室的实地工作人员可以视察有关地区。在没有发生重大变化的地区，可以按照经过的时间，根据上一次的人口普查调整估计数。

#### 4. 管理(人口普查小组领导)区划界

3. 27. 管理区为有效管理所有点查员提供了手段。点查区的划界完毕后, 设计管理地图一般就很简单。管理区通常由相邻的8至12个性质相同的点查区构成。划入同一管理区的点查区必须离得很近, 旅途不需花费很长时间, 并且大小基本上差不多。它们应被划入同一现场办公室的管辖范围, 现场办公室通常是按照行政管理单位划分的。

3. 28. 可以根据国家的大小设计其他级别的人口普查管理区。在较大的国家里, 这些管理区通常与省一级或区一级的统计办公室相当。

#### 5. 点查区的地理编码

3. 29. 向量格式的数字化地理数据库包括一系列点、线和多边形, 这些点、线和多边形构成一个整体。每一个地理特征, 即每一点、线或面, 都有一个惟一的标识符, 供系统内部使用。用户通常不能访问这个内部标识符, 更不能在外面对其进行修改。系统需要的是一个更有意义的标识符, 能把地理特征与其记录的属性联系在一起。对于点查区和行政单位, 这个连接是惟一的点查区或行政标识符, 所有与人口普查有关的地理区域主文件中都含有这个标识符。

3. 30. 如何输入这个标识符取决于具体的软件。可在进行数字化时添加这个标识符, 办法是在对地理特征进行数字化之前将其输入; 也可以在稍后时再添加, 办法是通过交互式方法选择特征, 通过菜单界面添加进去。对于多边形的特征, 有些系统要求用户添加一个包含在每一面积单位里的标记点。概念虽然很简单, 但编码所需的时间和资源可能相当可观。

3. 31. 事实上, 每个点查区都应有属于自己的唯一代码。这个代码用于数据处理, 以便对每个点查区的住户点查信息进行编制, 进行资料汇总, 供行政或统计区出版。这一数字代码还起到了连接汇总人口普查数据和保存在地理信息系统内的点查区边界数据库的作用。理想的编码方案应根据不同国家的具体情况决定。不过, 指定代码的规定务必清楚明了, 并且通过国家统计办公室内部, 特别是地图绘制/地理空间数据部门和管理核心数据的部门之间的合作进行设计。设计编码方案的最重要的原则是灵活性、扩展性、直观性, 并且与该国现行的其他编码方案的兼容性。统计办公室通常是该国编码方案的管理者, 而且还要在设计人口普查制图代码方面发挥关键作用。

3. 32. 分类编码方案通常可以使数字标识符一致而明确。按照这一方法, 行政管理体系中每一级别的地理单位都有自己的数字代码。通常在数位之间都留有余地, 以便将来插入该级新建区域的代码。换言之, 留下的空隙取决于如何添加单位。例如, 省一级单位可能会按照5、10、15的顺序依次类推地编码。较低一级的行政管理单位和点查区也采取类似方案。由于一个省所含的区通常多于一个国家所含的省的数目, 因此低级的编码系统所用的数位也更多。每一个最低级单位(点查区)的惟一标识符又与其所属行政管理单位的标识符相连。

3. 33. 例如, 一个国家可以采用下列编码方案:

图III.3  
一个普通的点查区编码方案

1	2	0	3	5	0	1	7	5	0	0	2	3
省		区			地区			点查区				

3. 34. 点查区代码1203501750023的含义是：第23号点查区位于第12省第35区第175号地点。这个惟一的代码以一个很长的整数或13个字节的字符串变量的形式保存在数据库中。显然，变量类型要与人口普查数据库和地理数据库中的类型相同。作为整数变量保存的好处在于，在任何数据库管理系统或地理信息系统软件包中，通过标准数据库的查询命令都可以轻而易举地选择出记录的子集。

3. 35. 另一方面，将代码作为字节变量保存可以改善代码的一致性，例如，通过先导零的使用。在这种情况下，代码更像是一个名字而不是一连串数字，这个名字由《美国信息交换标准代码》(ASCII)的字符组成。

3. 36. 如果行政管理单位和汇报单位部分等级，则要采用特殊的编码规定。无论在何种情况下，作为地理信息系统边界和人口普查数据表之间的链接手段，该行政管理单位标识符在规定和使用上要完全一致，这一点很重要。因此，人口普查官员应保留一个点查区和行政单位及其相关代码的主目录，并且应把主目录所做的任何变更输入地理信息系统和人口普查数据库。国家统计局办公室还可以考虑公布一个点查区清单，包括相关的编码和矩心(中心点)的经纬度，还可以包含已点查的人口数。

## 6. 人口普查数据库的组成

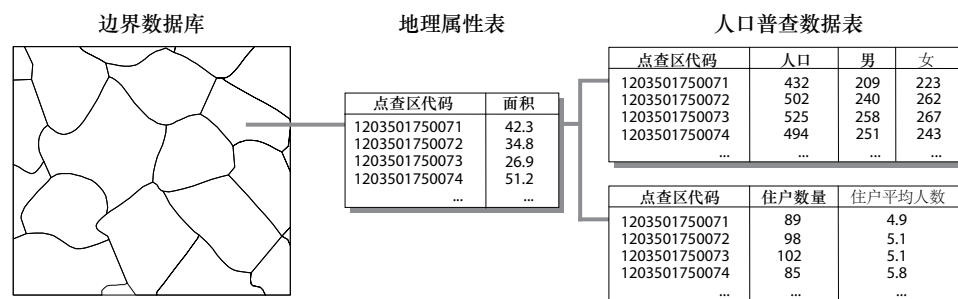
3. 37. 综合的人口普查地理数据库由点查区的数字化地图组成，在大多数情况下，由一系列底图层组成，在点查员最终使用的地图中为其提供背景和方位。基本数据层可能是各种道路、河流、建筑物或居民点。每个数据层都将被单独列入地理数据库。因此，举例来说，尽管道路和河流都用线表示，但它们不被存储在同一数字文件中。有必要提供书面说明，规定输出、大小和其他要求，以确保一致性。

3. 38. 开始进行数据输入和数据转换前，人口普查绘图人员应设计出即将制作的所有地理数据集的结构。结构定义应详细说明绘图人员必须遵循的所有惯例和准则，以确保最终产品输出的一致性。合理的规划和文件记录将避免在这一过程中发生混乱和随后出现不兼容现象。

3. 39. 第一步是思考点查员将使用的地图会是什么样的。例如，全数字点查区数据库可能由以下特征和属性因素组成(另见图III. 4)：

- 空间边界数据库，包括代表人口普查单位的面积特征(多边形)。
- 地理属性表。数据库文件以内部方式连接空间数据库并且载有每个多边形的记录。地理属性表载有每个普查单位的独特标识符，并且可能

图 III.4  
数字化空间普查数据库的组成



载有其他一些静态变量，也就是说不发生变化的变量，如以平方公里度量的单位面积。

- 普查数据表，载有非空间属性，即空间人口普查单位的普查指标。每一个这类文件均须载有人口普查单位的独特标识符，用以连接相应的多边形属性表记录。每个人口普查单位有一份记录。
- 另外，其他向量(点或面)特征，例如路标、公路、水道、学校、卫生设施或其他建筑物可能有助于实地工作人员在点查期间确定方向。在初步进行的实地调查或房屋列表期间记录的这类特征随后可能会证明对其他政府机构或非政府组织是有用的，可以节省时间和金钱。与其他数据用户合作在提高效率和效力方面有很多好处，在可行的情况下应开展这样的合作。

## 7. 点查区的设计与过去的人口普查保持一致

3. 40. 人口普查提供的是有关一个国家的人口数量和特征的概况。人口普查最重要的用途之一是分析特定时期内的人口组成变化。这一变化分析通常只在相当大的级别上进行，例如，在国家一级或者省一级上进行。然而，局部区域的变化同样重要，因为小区域的变化会影响地方规划决定。如果两次人口普查之间的点查单位可以相互兼容，地方一级的变化分析要简单得多。改动两次人口普查之间的地理基本资料所引发的问题，其严重性不亚于改动人口普查问卷上的问题。

3. 41. 在设计人口普查的地理因素时，人口普查办公室应当尽可能地保留上次人口普查所设定的边界。由于人口数量的增长，可能需要设定新的点查区。在这种情况下，从现有的点查区里划分新区总比改动边界要好。分析人员只要把新点查区的新的人口普查数据汇合在一起，使其与上次的点查资料保持一致就可以了。如果边界发生了变化，就需要通过更复杂的办法进行调整了。

3. 42. 绘制点查区轮廓图有利于进行改变分析，办法之一是对兼容性或等效文件进行编制，这些文件有时也被称为关系文件。这些文件列出了本次人

口普查中每个点查区的代码以及上次点查时的相应代码。如果点查单位发生分立或合并，它们同样能显示在这些文件中。

3. 43. 边界数据库和地理属性表紧密相连，实际上，它们代表一个数据集。在人口普查规划期间，将为各点查区编辑与基本人口普查有关的信息，如住宅单位和人口估计数以及文献资料等。这种有关人口普查单元的外部信息将被存储在类属数据库管理系统的单独数据表内。这些外部信息在数据表内可根据需要通过地理属性表内的通用标识符——点查区代码——连接边界数据。为了保证通过数据输入和制表方案生成的人口普查数据库匹配地理边界文件，人口普查绘图和数据处理部门必须开展密切合作。

3. 44. 一般来说，要为每一行政级别或已公布人口普查数据的统计区开发独立的数据库。如果任何级别的边界得到更新，载有这些边界的其他所有数据库当然也必须进行更新。最好的做法是全部更新最低加总级的主边界数据库（即点查区一级的数据库），并利用标准地理信息系统和数据库加总功能制作高一级的行政或统计单位数据库。

3. 45. 一些基本数据层可能比数字化人口普查点查区地图简单得多。例如，就某个道路数据库来说，也许仅收集了少数属性，即道路的名称或标识符（如果有用的话），还收集有地面类型和巷道数目。在这种情况下，不必将描述性属性信息存储在单独的表内。为了简便起见，所有属性均可列入地理属性表内。

3. 46. 在人口普查各周期之间及期间的某些阶段，应创建基准数据集。例如，应当制作具有独特性的国家人口普查地图数据库，并且对各项数据收集工作或相关的统计应用进行匹配。可以为需要提供数据的各统计报告单位编制单独加总的边界数据集。这些基准数据集应打上时间戳，永久保存。因此，从同一主数据库生成的基准数据集可以供2010年人口普查、2012年大型调查和2015年选举使用。

3. 47. 现有数字数据库，如另一政府机构制作的产品以及采用全球定位系统实地收集的坐标须输入人口普查地理数据库。全球定位系统坐标需从点位转换成表述道路或市区等的线性和多边形特征的线和界线。将属性代码附加到全部数据库特征后，可同时采用单独制作的数字地图创建整个地区的无缝数据库。完成的数据库——依绘图工作的范围而定——显示主要的物理特征、地标、基础设施、居民点及单个建筑物。根据这些信息，人口普查员能够以地理基准信息为背景采用交互方式划定点查区的边界。

3. 48. 第四章和附件二更详细地介绍了作为绘制点查区地图和评估点查区的数据来源的全球定位系统。关于使用全球定位系统和遥感数据来纠正地理数据库还提供了其他材料。

3. 49. 在整个数据开发过程同时进行的一项工作是，普查员须维护在数据库中描述的所有行政及点查区目录，包括名称、地理级别和位置基准。这个计算机化的目录是地理属性表，将连接完成的地理信息系统数据库。

3. 50. 图III. 5的流程图仅说明数据转换过程的诸如序列中的一个。在数据转换过程中，特别是点查区边界可以在若干点划界。例如，扫描和地理基准正确的航摄照片或者高分辨率的卫星图像提供了足够细节，操作员可以航摄照片为背景在屏幕上描绘点查区边界。点查区边界也可以用手工绘制在适当的纸图上，并与其他硬拷贝信息一起进行数字化处理。其他步骤也可以采用不同的顺序完成。例如，大多数地理信息系统软件包在数字化过程的开始阶段支持地理基准的确定，因此，随后的阶段不必采取其他步骤。

3. 51. 无论选择使用哪种过程，人口普查办公室都应进行试点研究，评估所采用的方法的可行性。评估主要涉及在小样本区域对方法进行验证。试点研究能够尽早发现问题，以便及时调整和改进技术和规程，或者在最坏的情况下予以废弃。试点研究形成的材料也有助于制定时间表和编制预算，因为这些材料能够更好地评估人员配备和设备需求以及完成各项工作所需的时间。

3. 52. 试点区应代表国家尽可能多的地区。换句话说，试点区的特征变化程度应当很大，包括农村及城市地区，居住方式典型化的地区、农业用地、浓密的植被区或制约实地数据收集的其他特征。

3. 53. 地理信息系统软件及设备销售商通常愿意协助进行试点研究，因为他们希望，如果自己的产品在试点研究期间经证明适用于人口普查绘图项目，产品的销售将为他们带来收益。销售商还将提供基准数据，这对于大量地图绘制和数据库存取等大容量应用来说是重要的。有些技术可以轻而易举地在国家的某个地区得到验证。例如，全球定位接收器的费用低廉，人口普查员能够评估实地数据技术。不过，小型试点试验场所的航摄数码照片费用过于高昂，无法获取。在这种情况下，可以使用情况相似的老式产品或航空摄影样片。

## C. 用于点查区划界的地理数据源

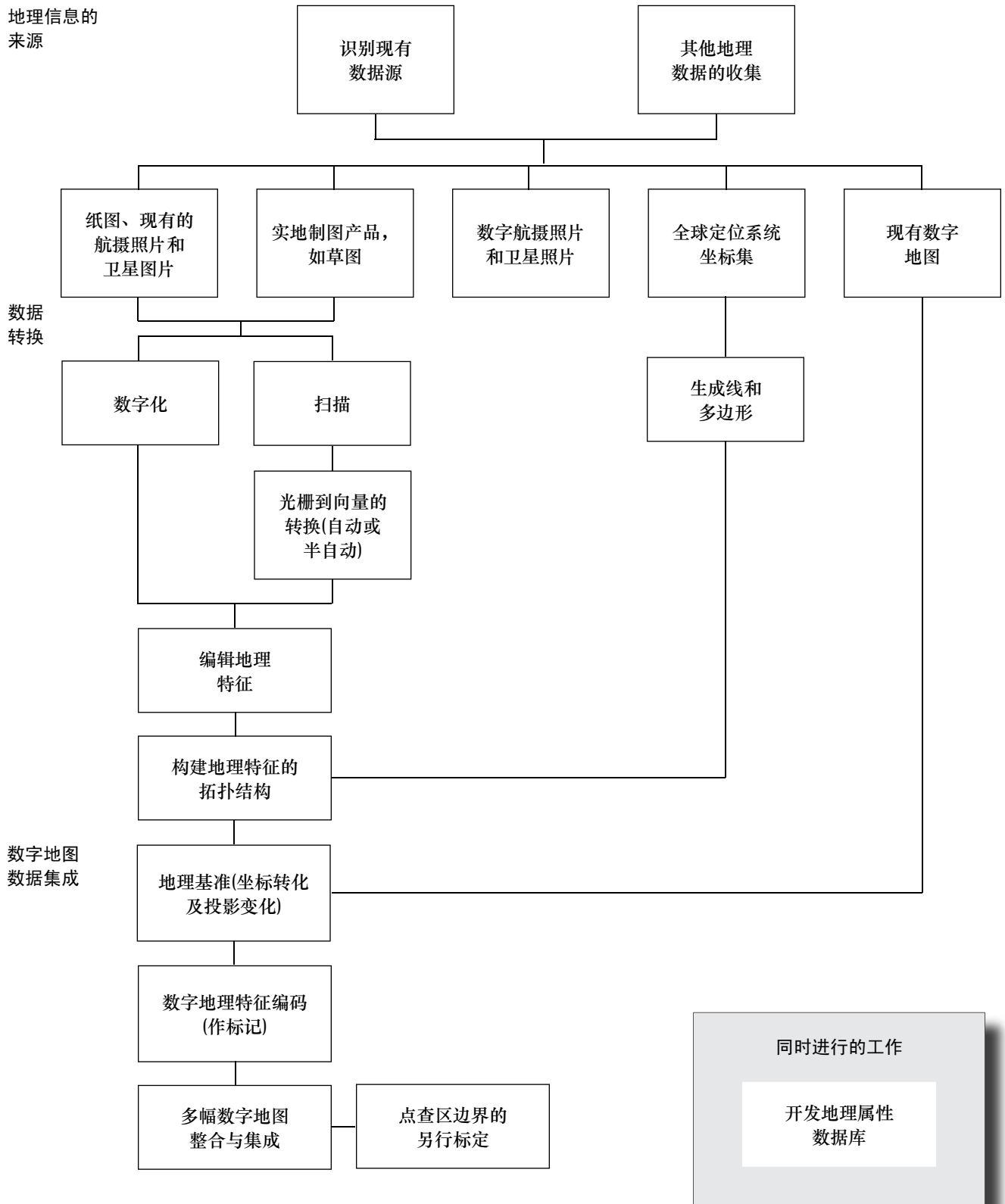
### 1. 所需地图的类型

3. 54. 几乎在所有情况下，人口普查制图方案均须参考硬拷贝地图来编制数字制图数据库，或更新现有的地理信息系统数据库。人口普查地理人员须获得国家的所有最新版地图，包括：

- (a) 国家概图，比例尺取决于国土面积的大小，通常在1:250 000和1:5 000 000之间。这些地图应标出重要居民区，城区的方位，重要道路、河流、湖泊、海拔及特殊参考点等主要自然特征。这些地图用于规划目的；
- (b) 大、中比例尺地形图。这种比例尺的地图的可用性因国家的不同而异。有些国家制作了完全覆盖国土面积的地图，比例尺为1:25 000或1:50 000，其他国家的最大的全图的比例尺仅为1:100 000或1:250 000；



图III.5  
人口普查地理数据库开发的各个阶段



- (c) 大比例尺城镇图，标有道路、城市街区、公园等。这些地图的比例尺从1:5 000到1:20 000不等，来源可能也不同，或许包括有年头的殖民地地图和城市规划；
- (d) 居民区各级行政单位的地图；
- (e) 专题地图，标有上一轮人口普查期间的人口分布情况，或对人口普查绘图有用的特征。

3. 55. 为了将这些地图编入地理信息系统数据库，最好应为它们编制综合文档。文档内容包括地理基准信息(如地图比例尺)、投影和地理数据、地图编制日期、编制机构以及图例大全。不过，即使地图的地理基准不正确，只要标出与人口普查绘图有关的信息，也是有用处的，尤其是在可以很容易地对其进行扫描并通过所谓的“平视数字化”(这已经变成了一个新定义)作为一种屏幕上可追踪的数据层纳入一个项目的时候。在这种情况下，附加信息产生的益处常常大于将这类数据合并到人口普查地理信息系统数据库所需的资源以及与这类产品有关的精度问题。

## 2. 现有资源库存

3. 56. 获得的所有地图都应按照人口普查方案的组织——即人口普查区——进行详细记录和整理。除了硬拷贝地图源之外，数字地图源可以不断地从许多渠道获得。当然，数字地图的优点是易于使用，适用于各种人口普查绘图目的。不过，并不能始终完全肯定这一点。如果缺乏文件记录和元数据，通常不可能确定正确的投影信息，数据质量难以得到评估(关于元数据开发的更多信息见下文F节)。

3. 57. 通过国家空间数据基础结构赞助的活动或者通过个人之间的接触，可同下述机构联系，看它们是否能够提供有用的硬拷贝或数字地图：

- (a) 国家地理机构/测绘机构。该机构常常是国家主管制图工作的主要机构，并且也许已经开始了地形图的数字化方案。不过，在一些国家，制图机构也许缺乏资源或法定授权来与国家统计办公室广泛开展合作；
- (b) 军事测绘机构。在有的国家，主要测绘单位隶属军方。军事测绘单位在航空摄影及遥感数据的判读方面常常拥有雄厚的实力；
- (c) 省、区及市政府。还可能包括城镇规划办公室。越来越多的地方政府机构利用地理空间技术管理有关交通、社会服务、公用事业及规划的信息；
- (d) 处理空间数据的各政府或私营组织，包括：
  - (i) 地质或水文勘查机构；
  - (ii) 环境保护机构；

- (iii) 交通机构;
  - (iv) 农村电气化管理机构;
  - (v) 公用事业和通信部门的公司;
  - (vi) 地籍部门;
- (e) 捐助活动。跨国或双边援助组织开展的项目级活动有时包括绘图活动。这类项目通常有办法购买并分析遥感数据或航摄照片,这对绘图机构非常有用。

### 3. 导入现有的数字化数据

3.58. 在大多数情况下,最容易的数字空间数据转换形式是直接导入数字化数据。地理信息系统行业已经从独立的表格(通常是DBASE(.dbf)格式的)转向了关系型数据库,例如Oracle或Microsoft Access,或通过个人或文件地理数据库。利用商业地理信息系统软件包的导入/导出功能,数据转换要依赖采用专用文件格式的数据的交换。

#### 方框III.1

#### 地理空间软件的选择标准: COTS、图像分析和FOSS选择

##### 1. 商业地理信息系统软件的选择

由于商业软件的供应商很多,国家统计局办公室应评估自己的业务目的并采纳相应的技术。软件互用性对于满足目前和未来的需要可能具有重要意义。关于首选的软件平台应咨询不同的机构和人员。知情的选择是最好的选择。

商业现货(COTS)软件可以分为提供光栅/向量集成的软件和主要进行图像分析的软件。

##### 2. 光栅/向量集成

软件包括环境系统研究所(ESRI)的ArcGIS、Intergraph的Geomedia、IDRISI、Maptitude、GRASS、Pitney Bowes的MapInfo、AutoCAD和Microstation。

环境系统研究所是市场领导者,提供广泛的格式支持、供不同用户使用的灵活的功能、广泛的分析工具、对各种数据类型的数据库管理、广泛的支持、培训和令人印象深刻的知识库。

GeoMedia支持许多数据类型,提供广泛的分析工具以及对主要的地理空间/CAD数据格式的直接访问。GeoMedia包含符合行业标准的关系型数据库,并且定期更新,同时提供全面支持和培训。

IDRISI是一种单一的打包产品,为定制提供开放代码。IDRISI提供基于光栅的复杂分析、光栅分析工具和通过CartaLinux进行的基于向量的广泛的数据输入和输出。

Maptitude操作CAD和地理信息系统,有输入/输出选项,提供地址匹配特征并且适合于随时可用的数据集。

地理资源分析支持系统(GRASS)提供光栅/向量能力,有一个新的拓扑2D/3D向量驱动器,并支持向量网络分析。GRASS是第一个UNIX地理信息系统,提供UNIX功能、开放代码以及一个全球用户库。

MapInfo (Pitney Bowes软件)提供制图功能,但只提供有限的地理信息系统功能。传统上,它使用Visual Basic和灵活的应用软件。它有一个全球数据库和一个强大的发布工具。

AutoCAD Map提供复杂的GIS/CAD集成, 以及网格、投影和拓扑支持。它提供广泛的数据库选择以及向量/光栅显示和分析。它以菜单为基础, 但价格不菲。

Microstation在美国和欧洲有大量用户。它以基于CAD的制图为特色, 但只有有限的分析功能, 以及有限的数据库格式集成和地理基准制定功能。它提供一个简单的Google Earth/Google SketchUp界面。

### 3. 图像分析

在人口普查制图活动中, 基于光栅的分析变得普遍起来。如今有很多图像分析软件能够集成向量数据, 以便进行精确的图像分析。由于图像已经变得不那么昂贵了, 处理起来更快, 并且与现有的向量数据集集成在一起, 更多的国家统计局办公室将购买图像分析许可。

主要平台包括Leica Geosystems Imagine、Geomatica (PCI Geomatics)、ENVI (ITT visuals)、Definiens Professional和Google Earth。

Leica Geosystems Imagine是基于光栅的软件, 设计目的是从图像中提取信息。它处理广泛收集的地理空间数据, 提供与空间数据库的客户端互动。Imagine提供空间数据库的拓扑编辑功能, 并提供向量数据层支持。

Geomatica (PCI Geomatics)处理广泛搜集的地理空间数据格式, 提供与空间数据库的客户端互动; 全面的地图生成功能; 用于属性查看、编辑、查询和分析的属性管理工具; 以及广泛的图像处理功能。

ENVI 4.5(ITT Solutions)将光栅图像与地理信息系统集成在一起。可以在图像数据上叠加向量数据层, 以便于对比光栅信息和向量信息。ENVI提供相对广泛的支持和一个广泛的知识库。此外, 它处理多种向量格式(包括ArcView图形文件、ARC/INFO互换、DXF、Microstation/Intergraph DGN文件、USGS DLG文件等)。其线性特征提取工具自动对种子点之间的一切进行数字化, 忠实地追随曲线、跳过缺口, 突然抓住向量, 如果想这么做的话。

Definiens可以进行极为复杂的、自动、用户定义的特征提取, 可以从任何种类的遥感图像中提取地理信息。它提供与ArcGIS服务器的数据管理连通, 可以向数据库载入和保存向量数据, 反之亦可, 可以同步更新不同位置的一个大型数据集。此外, 还可以更新大型向量数据集的一部分。对ArcGIS进行的扩展使得ArcCatalog可以把收集的地图定义为Definiens的工作空间, 使用户可以检查和编辑条目。

Google Earth最大的卖点是它的免费图像。不过, 存在与特征更新有关费用, 例如Google Earth Pro。Google Earth覆盖全球, 许多用户都可以访问。它对用户来说非常方便, 有一个大型知识库, 对于快速显示光栅和向量数据非常有用。

要考虑的问题包括软件的初始费用、维护和更新、LAN配置、培训需要、方便安装、维护、文件记录和手册、帮助热线和卖方支持、为支持打补丁的手段以及劳动力。

### 4. 用于桌面地图绘制的免费的开放源软件

除商业软件外, 还有免费的开放源(FOSS)地理信息系统软件。FOSS提供地理信息系统的低成本方式, 甚至无成本方式。FOSS软件可从因特网上免费下载, 并且提供与商业软件类似的功能。FOSS意味着用户可以获得应用程序的源代码, 就是说拥有程序设计方面的专门技术的国家统计局办公室可以定制软件以满足自己的特殊需要。FOSS软件正在变得更方便用户使用, 提供为特殊用途定制应用程序的功能。尝试使用FOSS软件的经验也许证明是宝贵的, 即使国家统计局办公室最终决定采用商业产品。FOSS桌面地图绘制软件传统上是编程人员或拥有信息技术经验的其他人使用的。由于用户基础扩大以及后来的产品开发, 这种情况已经发生了可喜的变化。软件变得更便于用户使用, 有在线培训和产品支持。FOSS软件提供“互用性”, 根据全球空间数据基础结构协会的定义, “互用性”是指“在各种功能单位之间进行交流、执行程序或传输数据的能力, 用户对这些单位的独特特征只需有少许了解或者不需要了解”。

FOSS软件包括以下这些：

Quantum GIS(<http://qgis.org>)。QGIS是最先进的光栅/向量FOSS软件，用户遍布六大洲，有一套快速演变的特征。QGIS提供GRASS，用于对向量特征进行数字化和编辑。一年一度的软件更新已经进行了多次，让用户总能使用最新产品。一个强大的用户团体对许多技术问题提供在线支持。

Thuban (<http://thuban.intevation.org>)。Thuban是FOSS软件的另一个例子，有很大的用户群以及在线的产品支持。Thuban通过Python执行，并且是多平台的。它既可以处理向量数据，也可以处理光栅数据，并且提供全面的地理信息系统功能，包括目标ID和注解、一个图例编辑器和归类、表格查询和连接、投影支持和多语言支持。

OpenEV(<http://openev.sourceforge.net>)。既是一个软件库，也是一个用于查看和分析光栅和向量地理空间数据的应用软件。OpenEV可以支持2D和3D显示，并且能够在中途进行重新投影。它提供强大的图像分析功能。OpenEV虚拟用户群利用讨论表和操作小贴士来接纳新用户。

3. 59. 所有软件系统都能连接其他格式，但例行程序的导入数量和功能因软件包的不同而异。由于软件开发商不愿公布其系统使用的确切文件格式，常常会发生一些问题。因此，在数据转换方面会出现困难。于是，竞争对手采用逆向工程设计方法来破解确切的文件格式，以便让用户导入外部文件。因此，导入例行程序有时不稳定，而且经常会丢失原始数据文件所具有的一部分信息。有时，最好采用另一种数据格式进行导入，不要试图直接导入另一软件包的交换文件。例如，Autocad绘图交换格式(DXF)得到了大多数地理信息系统软件包的支持，并且编制了详尽的记录文件。因此，其他商业软件包的DXF导出和导入功能通常非常可靠。

3. 60. 如果人口普查地理信息系统或者制图机构使用的是广泛使用的综合商业地理信息系统软件包，将会减少问题的发生。高端系统为许多交换格式提供导入功能的可能性更大。可能性更大的还有，其他数据编制人员能够以地理信息系统软件包的本机格式编制地理信息系统数据。导入能力的大小是选用地理信息系统软件的一个重要标准。另一种选择是使用第三方提供的转换软件包。

3. 61. 除了将数据文件从一种格式转换成另一种格式产生的问题以外，使用现有数字数据最常遇到的问题是元数据不足或短缺。如果没有元数据信息，就难以评估数字信息的质量。更糟糕的是，如果地理基准框架信息丢失，数据将不可能从外部数据集的坐标系转换为人口普查单位使用的坐标系。同样，如果代码手册或数据字典遗失，将难以判读地理信息系统数据集的属性表列入的地理及数据属性。如果数据的来源是外部，人口普查局应始终要求提供大量记录文件和元数据。

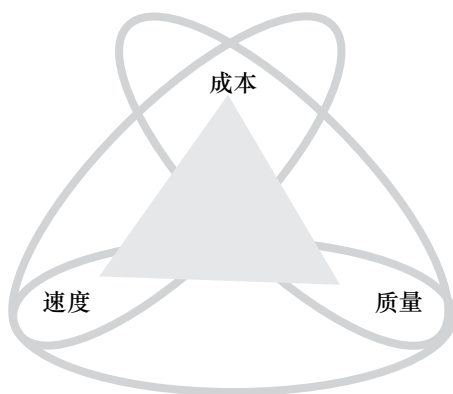
3. 62. 需要解决的其他问题有：定义和编码方案不同问题、不同制图基准系统的使用问题、空间比例尺不一致问题以及精确度标准各异问题，最后这个问题可导致两个数据库间应匹配的特征被置换。为了充分利用现有数字地图，解决这些问题可能需要进行大量的处理和编辑工作。

#### 4. 地理数据的转换：从模拟到数字

3. 63. 数字人口普查数据库的开发依据两个数据源：转换和结合可能采取硬拷贝或数字形式的现有地图产品，以及利用实地工作、航摄照片或卫星图像收集的其他数据。总的来说，“数据转换”这一术语用来指这些步骤。数据转换的指导原则是审慎规划。如果在这一过程的早期多加小心，那么在工作中就会减少问题的出现。

3. 64. 最佳数据转换策略取决于诸多因素，包括数据可用性以及时间和资源限制。项目成本、完成数据转换所需时间以及最终产品的质量之间总是存在着一种制约关系(图III. 6)。在这三项目标中，通常仅可能最大限度地实现其中的两项目标，而放弃第三项目标。例如，迅速创建一个高质量的数据库是可能的，但耗资巨大。编制有用数据的费用可以降到很低，但所需时间却很漫长。或者说，可以迅速开发数据库并降低开发费用，但最终产品的质量却很低劣。

图III.6  
数据转换过程的相互制约  
(依据Hohl, 1998年)



3. 65. 图III. 5概括了使完备的数字人口普查数据库得以开发的数据转换过程的基本步骤。对现有数字和硬拷贝源的调查结果是确定了数据间的差距。现有地图可能已过时，或者可用地形图的比例尺不能满足人口普查的要求。对于现有材料质量不达标地区而言，必须制定实地制图策略或其他一些数据收集方法。

3. 66. 人口普查所需地理特征的边界和点的方位(建筑物和村庄的方位、道路基础设施、河流以及点查区划界工作采用的其他信息)须根据已出版的纸图、草图、航摄照片或卫星图片采用数字方式标出。具体工作是通过鼠式光标对各种特征进行数字化跟踪，或者通过扫描把图像转换成向量形式完成的。尽管数字化和扫描技术在不断改进，这项工作仍然是数据转换过程最乏味的工作。捕获数据后的工作步骤是编辑，即构建地理信息系统数据库拓扑结构以及正确绘图投影中的所有坐标的基准(这个步骤有时可融入数字化工作)。

图 III.7  
进给式扫描仪的照片



资料来源：Ideal.com。

3. 67. 把硬拷贝地图上可见的地理特征转换成数字化的点、线、多边形和属性信息的过程，称为数据自动处理或数据转换。在许多地理信息系统项目中，这一步骤仍需要大量的资源，尤其是时间。

3. 68. 把硬拷贝地图或印制的航摄照片，或遥感图像的信息转换到一个数字化的地理信息系统数据库中涉及一系列步骤。虽然步骤顺序可能有所不同，但所要进行的程序大同小异。首先要在地图上选择点状或线状的地理特征，将其转换成计算机内的数字化坐标值，通常还要进行大量的编辑，处理存留的误差和疏漏。之后，还要把数字化仪或扫描仪所使用的坐标的长度单位转换成原地图投影法采用的真实世界的相应坐标值。某些系统在数字化之前，可以确定地图投影方法。如果是这种情况，那么，在数字化的过程中，坐标值就会自动随之转换。当然，最终结果都是相同的。

3. 69. 下一个步骤是，为数字化地理特征加上统一代码。例如，代表道路的线段需要一个代表路面状况的编码(土路、单车道、双车道公路等)，或者赋予一个惟一的编号，以便与其他有关信息连接，例如，与街道名称表连接。在更高端地理信息系统软件包中，这一步骤之后，还要建立数据库，亦称建立拓扑关系。在这一步骤中，地理信息系统确定数据库内的各个地理特征之间的关系。例如，对于一个道路数据库来说，系统要确定两条或多条道路的交叉点，并在交叉点上建立一个节点。对于多边形数据而言，系统还要确定哪些线段对各多边形的边界进行界定。数字化数据库完成并检验无误后，最后一步是添加附加属性。有些信息可以永久性地连接到数据库上，而有些关于每个数据库特征的附加信息则可以存成单独的文件，这些文件都根据需要与地理数据库连接。

3. 70. 把硬拷贝地图的信息转换为数字化数据有两种主要的方法，一是扫描，二是数字化处理。扫描是一种自动把地图转换为一种数字化光栅图像的

过程，以后这种光栅图像可以转换成数字化线段。而数字化处理是利用光标或鼠标对地图上所有需要的点或线状特征进行跟踪。在屏幕上进行数字化的技术被用来对扫描的地图或图像提取新的地图层，也可以从一页页做了标识的地图上对内容进行数字化。下面将详细介绍这两种方法。

#### (a) 扫描

3.71. 对许多数据输入工作来说，扫描无疑已绕过数字化成为空间数据输入的主要方法，主要原因是通过大版面进给式扫描仪和互动式向量化软件，能够自动完成一些乏味的数据输入步骤。扫描仪有多种类型，但工作方式基本相同。将图面朝下放置在扫描平面上，光线斜照在图上。一个感光装置将每个小格或像素反射的光强记录在非常精密的网格上。在灰度模式中，光强直接转换成数值，例如，转换成0(黑)到255(白)之间的一个数。在双值模式中，以光强阈值为准，将光强转换成白或黑(0或1)单元值。在彩色扫描仪中，光感装置分成三部分，分别感应红、绿、蓝三种颜色。组合在一起时，三种颜色信号的相对强度就决定了像素的颜色。扫描的结果产生出原图的光栅图像，它能够储存在一个标准的图像格式里，如GIF或TIFF格式。在为图像加注地理基准信息之后——包括以实际使用的单位规定图像角坐标和像素的大小——图像就可以在许多地理信息系统软件包中以现有的向量数据的背景出现。而从图像中抽取地理特征的工作通常由人工完成或自动完成，并转换成向量数据。

#### 3.72. 常用的扫描仪有三种基本类型：

- **进给式扫描仪**是目前大型地理信息系统应用使用得最多的扫描仪。这种设备的传感系统静止不动，而地图在传感器阵列上横向移动。其精确度低于滚筒式扫描仪，因为与扫描头的运动相比，图像进给的控制较难做到十分精确。但对地理信息系统应用来说，其精确度通常已经足够，而成本却较低，图像扫描一般不超过5分钟。需要提请注意的一点是：进给式扫描仪的滚轮可能会对较陈旧的或易破碎的文件造成损坏(见图III.7)。
- **台式扫描仪**是许多办公室常用的设备。这类设备版面较小，因此，较大的图必须分成几部分扫描，然后由计算机拼合。文件面朝下放置在玻璃板上，摄像机和光源在玻璃板下顺着文件移动。台式扫描仪的优越性在于价格低廉，便于安装和维修。它们很适用于扫描文本文件(如数据表格)，然后再用光学字符识别软件解读。它们也能将小的图标和地图等输入到计算机中。这类设备较不适合于大型地图的转换工作，在这样的转换中，会有许多大版面的地形和专题地图需要扫描。将这些地图分割成小部分进行扫描，然后再在计算机里拼合，是一件耗时费力的事情，并且很可能因此产生大量错误。
- **滚筒式扫描仪**价格比较昂贵，适于精度要求非常高的专业用途(如摄影制图或医学方面的应用)。将地图固定在一个旋转鼓轮上。然后，一个传感系统顺着地图移动，记下每个像素的光强或颜色。滚筒式扫描仪



虽然精度很高，但它们过于昂贵且速度相当缓慢。一次扫描可能需要15至20分钟。

3.73. 操作人员对扫描仪进行的设定对输出图像的特征有巨大影响。选定最佳参数需要经过一定试验，因为这取决于扫描仪的选项、被扫描的底图或照片的特征以及预期的后续处理步骤。最重要的参数如下：

- **扫描模式。**双值或“线条”模式对单色图或草图是合适的，对色分离的所有特征基本属于同一类型的图也同样适合。灰度模式能在地图上保留变量，并可借助后续图像处理，在图标或图像处理系统中抽取仅具有一定反射值的特征。如果用彩色模式扫描地图，这就会更加容易，例如，地图上所有的绿色表示的形貌可以用几个简单的指令抽取出来。
- **图像分辨率。**用每英寸点数(dpi)表示。一般扫描分辨率在100到400 dpi之间(航摄照片通常使用专用扫描仪进行扫描，分辨率较高)。较高的扫描分辨率能够保存较多的原始地图的细节，在向量化的地理信息系统数据集中可得到较光滑的线条。但所得的图像尺寸较大，需要更多的内存和磁盘空间；扫描分辨率增加一倍，图像尺寸要增加四倍。如何选择分辨率取决于源文件的性质、现有硬件和最终图像的计用途。
- **亮度、对比度和阈值。**这些参数决定了最终图像的外观。亮度决定整体图像的明暗。对比度用于决定如何保存灰度和细微色调。对比度高使图像显得更加鲜明，但也可能因此丧失一些变量和细节。阈值这个参数用于双值模式中，决定原始文件中的灰值如何转换为黑白像素。有的扫描是为了产生对视觉有吸引力并能准确表达源文件的图像，有的则是为了进行后续向量化处理，扫描的目的不同，参数选择也就不同。对后一种目的来说，较高的对比度或亮度可以突出地图中的特征，因此有助于以后往向量格式转换。
- **伽马校正。**如果图像像素值比较规则地分布在全灰度范围内，亮度和对比度就能比较好地进行控制处理。但情况往往并非如此。例如，图像可能由非常亮和非常暗的区域组成。伽马校正这项技术就是考虑到图像中灰度值的分布，从而自动进行调整，使区域增量或变暗，或将单元值扩展到更广的灰度值范围去。这项技术常常能够帮助保存图像中的细微变化。

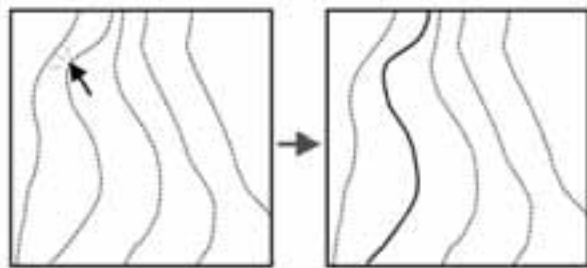
3.74. 扫描源文件仅仅是第一步，也是相当简单的一步。由于转换过程的最终结果是点和线组成的数字地理数据库，在光栅图像上包含的扫描信息需要转换成坐标信息。这个过程叫做“光栅向量转换”。直到最近，这个步骤还是扫描过程中的一个薄弱环节，这也是数字化一直经常被优先选为数据输入方式的原因。软件开发、图形识别技术和处理速度方面的最新成果使得许多人把扫描作为其首选的数据输入方式。

3. 75. 光栅向量转换的操作有自动、半自动和人工手动三种模式。在自动模式中，专门的软件将光栅图像的所有线条都自动转换为坐标序列。由于地图上的粗线到了光栅图像上就成了有几个像素宽的线条，自动的光栅向量处理过程总是从线条细化算法开始。下一步是决定每一个给线条定义的像素的坐标，接着，很可能就是除去多余的坐标——即用较少的坐标来表示直线。转换软件常常允许用户指定容差级。例如，只含有一个或几个像素的地方可能实际表示的只是源地图上的一个污点，因此可以自动消除。还有，如果图像是使用彩色扫描仪扫描而成的，光栅向量软件常常允许用户指定配给颜色的线条代码。这对于抽取不同类型的特征并入不同的地理信息系统数据层是有用的。例如，在源地图上，河流可能用蓝色表示，道路被画成黑色，而行政区划的界线则为红色。

3. 76. 在半自动模式中，操作员点击每条需要转换的线(见图III. 8)。此系统就循着这条线找到最近的交叉点，将它转换成一个向量表达式。这样做的好处是：操作者在地图上只需选择一组特征子集，例如，只选择所有的道路，而不要河流。最后，谈谈手动模式。扫描所得的光栅图像仅仅用作计算机屏幕上的底图。操作者利用鼠标随着扫描图像上的特征建立坐标，这和以前提到的平视数字化方法类似。

图III.8

#### 半自动向量化处理

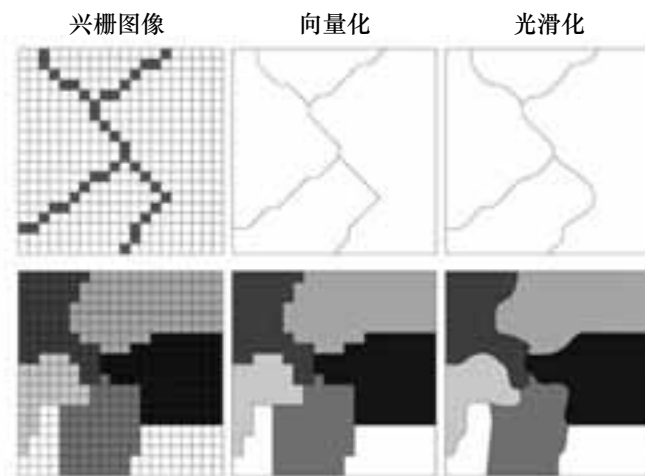


3. 77. 如果线性特征或区域特征是自动由分辨率相对较低的光栅图像转换成向量格式的，所得线形可能显得生硬。通用的办法是利用地理信息系统软件包中提供的样条功能或泛化功能使向量数据变得光滑。图III. 9显示了线形和多边形数据集两个例子。

#### (i) 其他考虑

3. 78. 在计划一个基于图像扫描的数据转换项目时，有许多要点需要考虑。在扫描前妥善准备地图底本有助于显著改进输出质量。地图应该平整干净。地图上的任何胶带残余物都应除去，否则它们会在扫描仪面板上留下痕迹。地图上模糊的地理特征应用钢笔或其他标记笔勾描清晰。同样，操作员能回填扫描的线形符号，填补有交叉阴影线的多边形，形成密实的线条和填补部

图III.9  
扫描图像数据的向量化和光滑化



分，便于自动向量化的进行。另外，这些改变也可以在向量化前在扫描的图像上操作。任何基于光栅的图形软件包都可用来做这项工作。然而，靠手工实现这些改变往往更加容易。修改时应该用水质标记笔或蜡铅笔，因为油质标记笔可能会对扫描仪的玻璃面板造成损坏，而石墨铅笔标示的记号会反光，使记号隐而不现。对于照片来说，布纹相纸比光面相纸会产生更好的效果。

3.79. 当转换较为复杂、特征繁多的地图(如地形图)或质量差的地图时，经常增加一个补充步骤。对于这些地图数据源，如果首先将所有需要的地图特征在聚酯薄膜一类的透明介质上描一遍，就可以提高精确度和减轻后续处理的劳动强度。虽然它增加了操作员的工作负荷，但最后往往证明，描图能够减少编辑和改错的时间，因此成为加快速度的手段。经过描绘的源文件扫描后比较清晰，而且仅包含真正需要的特征。大多数大规模专业扫描作业时经常需要这道工序。如果能取得源地图出版物的原始分色图，即可免去描图工作。制作全国系列地形图时常常能得到这些分色图。每张分色图只包含所印地图的一个特征子集，这就大大简化了将不同特征分离到不同数据层的工作。

3.80. 尽管采取了这些准备步骤，在启动向量化程序前，扫描所得的图像可能仍然需要做进一步的处理。这些处理可能包括图像进一步增强，如锐化或加大对比度，还有，去除斑点或在像素级进行交互改变。一个面向光栅的图形软件包或向量化软件本身都会提供所需的功能。

3.81. 支持光栅数据的地理信息系统软件包能提供光栅向量的转换程序。但是，这些程序大多用于光栅地理信息系统和向量地理信息系统数据之间的转换，而不是将复杂的、扫描所得的图像转换为纯向量特征。对于大规模向量化项目来说，使用专用软件包较为合适。目前有许多独立的商品化和非商品化的光栅向量软件包(如Vextractor、AbleVector和PTracer)以及软件的扩展(如ESRI ArcScan)问世。就这些产品而言，可供选用的功能各不相同。其中一些

具有纠正扫描图像倾斜的功能，或识别地图注释字符的光学字符识别功能，将这些注释存储起来，作为所得向量特征的属性。各种产品的价格也相差甚远。因此，数据转换工作人员应该将数据转换任务的需求与各类产品提供的功能进行认真比较。

#### (ii) 扫描的利弊

##### 3.82. 扫描的优点如下：

- 扫描的地图可用作向量信息的图像底图。例如，利用平视数字化，扫描的地形图可以和数字化的点查区边界合起来用于制作点查地图。
- 通过光栅向量转换软件，可以相对容易地对清晰的底图和原始分色进行向量化处理。
- 小版面的扫描仪价格相对低廉，并能提供快速的数据采集功能。

扫描的缺点如下：

- 用小版面的扫描仪转换大地图，需要通过枯燥乏味的劳动将各个部分拼合在一起。
- 虽然大版面和高效率扫描仪十分昂贵，但如果用于大比例地图的扫描及向量化从而得到数字化地理数据库，其费用还是合理的。
- 尽管向量化软件的开发近年来取得了进步，但仍可能需要进行大量手工编辑和属性标注工作。
- 扫描大量硬拷贝地图对于在许多台式计算机系统中存储文件来说是一个挑战。考虑对所有点查区地图进行扫描的国家统计办公室应考虑投资于一个带备份系统的单独的硬盘驱动器以适应大量的生成文件。

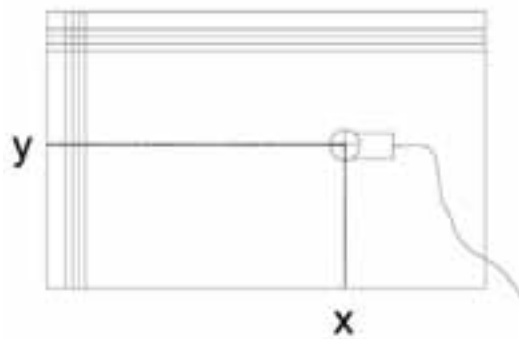
#### (b) 数字化处理

3.83. 对于空间数据的自动化处理来说，手工数字化处理在相当一段时间内是最为普通的方法。手工数字化处理需要一张数码化表，其大小从30x30cm到120x180cm不等。大型的数码化表有助于加快大型地图的数字化进程。用小块数码化表做大型地图的数字化工作必须将地图分成几块，完成后再拼接到一起。在数字化处理过程中，通过定标带把地图固定到数码化表上。比较理想的地图应该是平展的，既不要卷曲也不要折叠。纸张通常都会收缩，特别是在潮湿的条件下，这种收缩将导致纸张变形，最终使数字化地图数据库失真。

3.84. 数字化处理的第一步是在地图上准确地定出几个“控制点”（通常至少四个）。这些控制点有两个目的。一是如果一幅大型地图分几个阶段进行数字化处理，那么它必须在数码化表上移动多次，有了控制点就可以重新对准。二是要在底图投影系统中选择控制点的已知真实坐标。理想的控制点应选择在经纬线的交叉点上，在多幅地形图上均可看到。在对点或线形特征进行数字化处理之前或之后需要确定地理基准，在这个步骤中，要把数码化表上测量到的厘米或英寸转换成真实世界中的地图投影的坐标单位，通常是米或英尺。

在选定控制点之后，操作人员运用一个可与数码化表通信的光标器，在地图上跟踪线形特征。数码化表上有一套电路网格(它的部分形状见图III.10)。电路网格形成电磁场。光标器有一个金属线圈。数码化表与光标器之间发生相互作用，一个是发射器，另外一个接收器。因此光标器能够确定与其最近的x方向的电线和y方向的电线。高精度的准确位置可通过插值法获得。经过数字化处理的特征可立即送到计算机屏幕。操作人员可以看到已经捕捉到哪些边界，是否出现了很大误差。

图III.10  
数码化表



3. 85. 坐标值是按照点、距离或流线的模式记录的。在点模式中，线条的方向每改变一次，操作人员按一下光标器上的按钮。对于曲线来说，所记录的坐标点的数目决定了出现在地理信息系统数据库中的一条曲线的光滑程度。在距离模式中，坐标值是根据光标移动的规定距离自动记录的。而在流线模式中，坐标值是按照预先规定的时间间隔，在光标移动过程中自动记录。在距离和流线模式中，对于带有很多曲线的复杂线段存在一定危险，所记录的坐标点可能过于稀少。反之，对于长直线则可能产生出很多冗余点。一般来说，有经验的操作人员愿意使用点模式，自行确定坐标点的疏密程度。

3. 86. 对于操作人员来说，数字化过程往往是一件既繁琐又令人疲倦的工作。因此，除了安排训练有素的操作人员之外，还要提供良好的操作环境，包括一台符合人体工程设计的合适的数字化仪。可以辅助操作人员工作的性能一致的地理信息系统软件宏指令以及质量控制程序有助于减少数字化过程中的差错，缩短后期编辑的时间。

3. 87. 在进行数字化处理期间，操作人员可以给所捕捉的每一条线或每一个点指定一个特征编码。例如，对于不同级别的行政管理区可以这样编码：用一代表省级边界，用三代表区级边界。在某些具有拓扑结构的地理信息系统中，用户必须为每一个数字化的多边形增加一个标记点，这个过程可以在数字化处理时手工完成，或在建造拓扑结构之前自动完成。这个标记点的作用是，在多边形和含有多边形相关数据的地理属性表之间建立连接(见附件一)。

3. 88. 还有一种不用数码化表的数据输入方式，有时称之为平视数字化处理。平视数字化处理如今有两种含义。采用老方法时，操作人员对一幅透明

地图上的特征进行跟踪，而该地图置于计算机屏幕上。通过鼠标在地理信息系统数据输入模块或干脆在一个支持地理信息系统兼容格式的绘图软件包中，对线条和点进行数字化处理(见图III. 11)。采用平视数字化处理的新方法时，以数字化的方式利用一张扫描的地图图像，以便将轮廓描绘进地理信息系统层中。操作人员使用扫描地图、航摄照片或卫星图像作为背景。对图像已经进行过地理基准工作，也就是说图像已经转换为具有与地理信息系统项目中其他数据层相同的投影和现实世界的坐标系的一种格式。分析人员通过使用控制点和在剩余的数据层以及在现实世界中将图像“限定”在已知的位置来转换图像。好的控制点包括街道交叉点和陆标。分析人员随后借助鼠标从扫描的图像上跟踪地理特征，在这一过程中建立一个新的数据层。

#### 数字化的利弊

##### 3. 89. 数字化的优点如下：

- 数字化容易掌握，因此无需昂贵的熟练劳动力。
- 数字化过程中可随时添加属性信息。
- 通过人工数字化处理，可获得高度的准确性；即与原图相比，一般不会丧失准确度。

##### 3. 90. 数字化的缺点如下：

- 数字化单调乏味，很可能导致操作者过度疲劳，造成质量问题，因此可能需要大量的后续工作。

图III.11

平视数字化



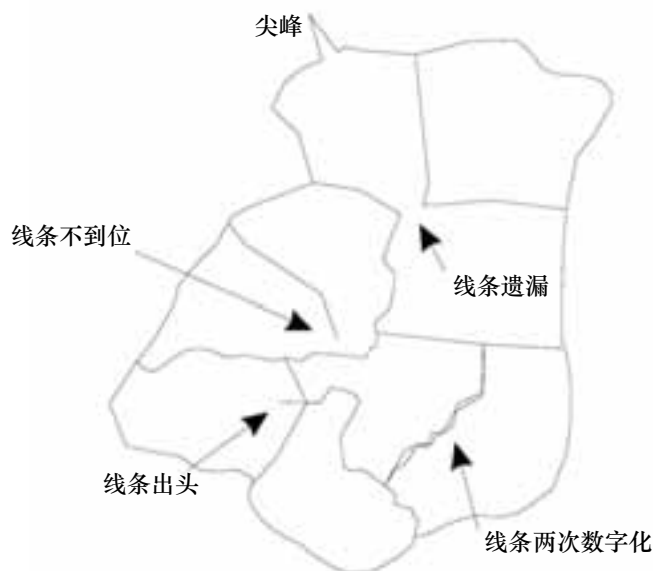
- 人工数字化处理速度比较慢。大规模的数据转换项目可能需要为数众多的操作人员和数码化表格。
- 与使用全球定位系统或空中摄影所收集的基本数据不同，数字化地图的准确度必然受原始资料质量的制约。

### (c) 编辑

3.91. 将地理信息从模拟形式转换为数字形式的目的是准确表达原始地图的数据。这就意味着地图上的所有连接线在数字化数据库里也必须相连。既不可遗漏任何特征，也不可重复任何线条。手工数字化处理容易出错。最常见的错误见图III.12。当发生线条断开和缺失的情况时，最终的影响是在实际存在两个区域时，只建立了一个多边形。在两次对线条进行数字化时，最终的影响是另外建立了一个或多个多边形，而实际上一个多边形都不存在。在光栅向量数据转换后，断开的线段需要用手工连接。例如，当在粗线表示的主干道上有细线符号画的小路或河流跨越时，就会发生线条断开的情况。如果将次要道路或河流抽取出来形成单独的地图层，在和主干道交叉处，道路网络上就会出现空白。

3.92. 图III.12所示的常见的数字化错误中，有一些是可以避免的，办法是利用用户定义的数字化软件中的所谓“链接公差”。例如，用户可以指定，一根线的所有端点与另一条线相距小于1毫米时，应自动与后者连接。一根线两次数字化时形成的细碎多边形也可自动消除。然而，用这个方法只能解决一部分问题。经过认真比较原始地图和数字化地图后，由手工改正数字化错误仍然是数据转换工作中的一个必要环节。

图III.12  
几种常见的数字化错误



#### (d) 建立和维持拓扑结构

3. 93. 转换地理数据，使点变成多边形的节点，这需要对其进行定义，让它们知道相对于其他对象而言它们是什么。这阐释了“拓扑”的概念，对拓扑的定义是研究几何图形不会因失真而改变的特性。拓扑正确的数据与没有拓扑结构的图表对象(常常被称为“意大利面条”)不同。有拓扑结构的对象知道自己在绝对空间中的位置，也知道它们最近的邻居。建立数字地图的拓扑结构有利于编辑工作。例如，这可以使用户鉴别一些问题，如不完全闭合的多边形等。特征拓扑可说明相连或相邻的地理特征之间的空间关系，如在交叉点相连的道路等(参见有关地理信息系统的附件一)。根据拓扑学建立的地理信息系统数据库包含对这些空间关系的鉴别和它们在数据库中的描述。实际的操作方法在各软件中有专门规定。存储拓扑信息可简化分析，因为许多地理信息系统操作并不真正需要坐标信息，而只是以拓扑为基础。例如，一个地区的邻接区就可以用一个数据库表格来确定，表上为每根线列出了位于它的左右两侧的多边形(详见附件，A1. 5段)。

3. 94. 用户一般必须考虑地理信息系统如何存储拓扑信息。假如数字数据库是经过清理的，也就是说，所有线条都已链接，所有多边形都已正确鉴别，那么，地理信息系统的一个功能可用来建立拓扑结构和创建所有需要的内部数据文件。成功使用这一功能的前提是，地图数据库不得包含任何错误。建立拓扑结构因此还能够起到检验数据库是否完整的作用。

#### (e) 数字地图的集成

3. 95. 一个人口普查制图项目应该利用一切适当的地理空间数据源。这些数据源很可能按照不同的地图比例和制图投影存储在不同的格式里。如果目的是要制作一个完整无缝的人口普查数字地图数据库，把这些杂乱的数据源集成在一起就需要大量的有关地理信息系统数据集成方法的知识。下文将探讨便于数字地图数据集成的一些最重要的方法。

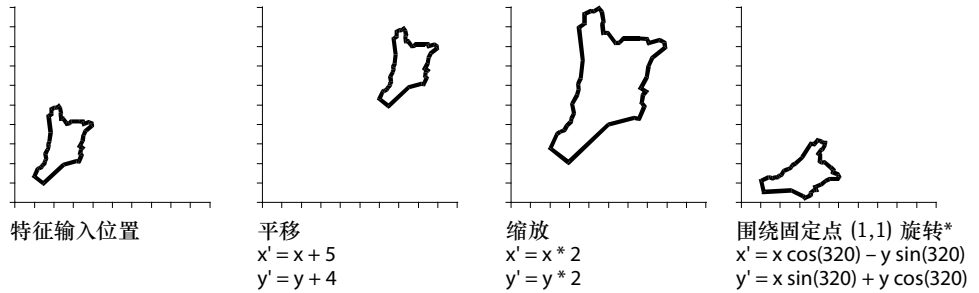
#### (f) 制定地理基准

3. 96. 制定地理基准是从已知参照系，例如经纬度的坐标转换为一个光栅(图像)地图的页面坐标的过程。几乎所有的地理信息系统软件包都提供制定地理基准所必需的功能。用户需要指定几个已知其真实世界坐标的控制点。根据以数字化单位表示的输入坐标数据和真实世界输出坐标，系统则对实现下列变换的一组变换参数进行计算(见图III. 13)：

- **平移**。地理特征被移至一个新的位置，办法很简单：给x和y坐标加(或减)一个常数值。x和y的偏置量往往不同。
- **缩放**。将x和y坐标各自乘以一个因子，即可将特征放大或缩小。缩放比例通常相对于坐标系的原点确定。



图III.13  
平移、缩放、旋转



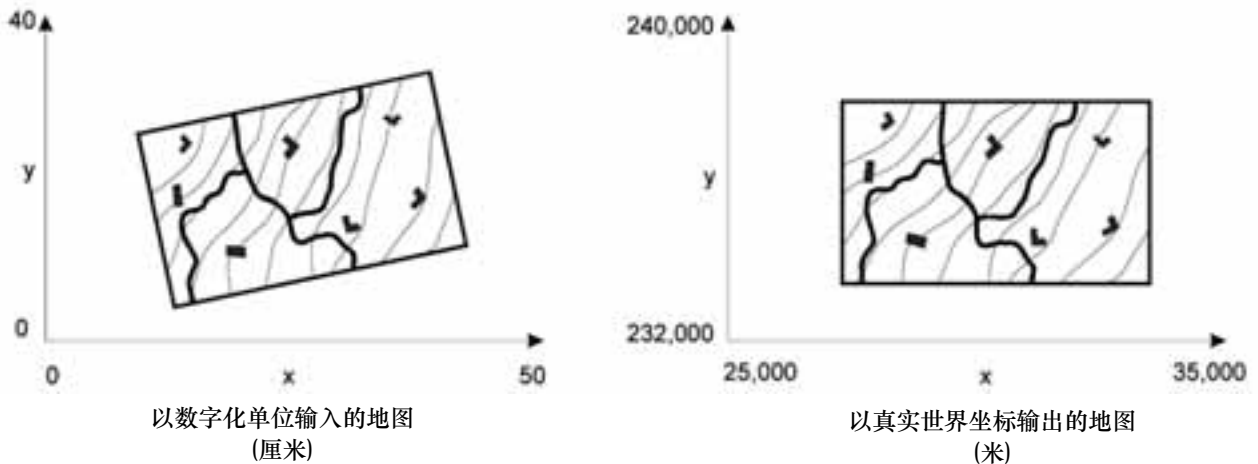
平移需要在围绕原点旋转之前和之后进行。  
旋转为正的逆时针方向。

- **旋转。**地理特征按一个给定角围绕坐标系的原点旋转。旋转能保证所得数字地图正确定位，即使地图纸在数字化板上没有对正也不会发生问题。

3. 97. 注意，数字化特征的形状在上述变换中没有发生改变，而将图像进行投影时形状也许会改变。只有对象的相对尺寸和方位经过了修改。在正确的平移、缩放和旋转参数计算出来以后，系统就将这些参数应用到数据库中的所有点和线的坐标上。输出的是一张看起来非常相似的地图，但现在使用的是原始底图的正确坐标系(见图III. 13)。重要的是尽量降低操作中出现的错误。系统通常为每个点提供转换参数估算中的错误信息，这将有助于查找出在指定控制点的真实世界坐标中出现的错误。更详细的技术细节见附件二。

3. 98. 如果不知道纸张形式的源地图的地图投影和坐标系，就会产生严重问题。遗憾的是，这类问题还经常发生，因为许多纸张式的地图，尤其是专题地图，并不包含这些信息。如果发生此类情况，有两个方法可以选择：尝试

图III.14  
以数字化单位表示的地图；以真实世界坐标表示的地图



通过大量可能的地图投影进行比较(应优先尝试国家地图规划中使用的标准投影),或使用所谓的橡胶板法。

3. 99. 橡胶板法需要大量均匀分布在地图上的控制点。有时,一张国家和行政区划的数字地图,或者任何其他同样存在于数字化地图里的清晰界定的点,都可以用来寻找相应点之间的联系。然后,系统就用输入和输出的坐标来计算更高阶的多项式变换。橡胶板法导致的误差一般都相当大,因此,只要有可能,就要避免使用这个方法。然而,在有些情况下,输入地图与任何恰当定义的投影明显不吻合,橡胶板法就成了利用现有地理信息的可行方法。在制作人口普查地图时,为人工绘制的草图制定地理基准就是一个很好的例子。附件二的A2. 6段列举了一个制定地理基准的实例,对转换过程进行说明,例如,将一张数字化地图转换为正确制定了地理基准的数字数据库的过程。

3. 100. 对利用全球定位系统收集的同一面积或坐标确定了地理基准的数据层将不与数字化地图兼容,因为它们是以真实世界坐标系为基准的。因此,数字化的点和线的坐标需要从数字化单位转换为以米或英尺衡量的真实世界地图坐标(另见附件二)。正如上文提到的,在大多数系统中都可以进行这个步骤,或者在数字化开始时,或者在空间数据自动化完成后。

#### (g) 投影和数据的变化

3. 101. 从不同的地图制作成的数字化地图,如果需要组合成无缝的数据库时,就必须改变投影。在转换数字地图特征坐标的同时不改变它们的形状,这样的变换过程关系到投影变化的问题。当从一个投影转换成另一个投影时,地图特征的形状和失真确实发生了变化,尽管这些变化在大比例图中几乎难以察觉。

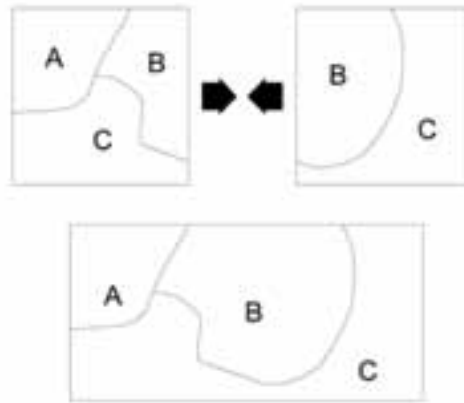
3. 102. 不同比例的地图有时使用不同的投影。还有一种情况是:制图机构可能已经更换了在该国用于地图绘制的标准投影,使旧地图可以使用与新近修订的地图不同的投影。同样,制图机构也可能已经修改了地理数据,建立了为该图制图用的基准框架,结果,比如说,旧地形图的坐标系就与新地图产生了细小的差别。

3. 103. 附件二对投影和地理数据进行了更详尽的探讨。人口普查制图机构应该配备训练有素的人员或者来自外部机构的专家,向他们咨询调整投影的最佳方法及有关问题,以生产统一的国家人口普查地图的底本。改变投影的实际操作所需的劳动量相对较少,因为所有商品化地理信息系统都提供所需的投影变化功能。

#### (h) 分立地图块的集成

3. 104. 数字化制图项目的目标是为一个大区域或整个国家制作无缝数据库。当制图比例为中到大时(如1:250 000或更大),底图的信息将包含在分立的地形图块上。这些图块将分别进行数字化处理,所得的数字化地图块将在地理信息系统中集成到一起(见图III. 15)。

图 III.15  
连接相邻的数字化地图块

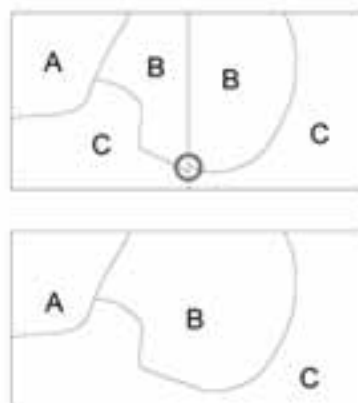


3. 105. 这一过程通常很简单，但图与图之间的衔接很难百分之百地保持完美。一些跨越两张图的特征(如道路或边界)在地图的边界上可能发生错位(见图III. 16)。错误可能是在数字化过程中产生的，或者在源地图上甚至就已经实际存在了。例如，相邻的地图可能是不同时间制作的，因此，较新的特征，如新的道路等，就没有跨过边界继续画下去，或者用了不同的符号表示。

3. 106. 假如没有按所需比例制作的完整覆盖全国的地图，因此需要拼合具有不同特征密度的不同比例的地图，这时问题就显得特别严重了。当拼合城乡结合部的地图时，常会遇到这个问题，因为大比例的城市地图需要和比例较小的农村地图配接。由于制图泛化方法的差异，在比例较小的地图上，某些特征可能显示，也可能不显示，或者它们的象征符号在两个地图系列中可能不一样。拼合这类地图需要认真判断和丰富的经验。在混合比例时保持谨慎是明智的。

3. 107. 修补这些错误的过程叫做边缘配接。通常用手工进行，涉及繁重的编辑工作。如果错位不太大，跨图纸的特征又能相容，则可利用某些地理信息系统软件包提供的自动边缘配接功能来连接这些特征。

图 III.16  
连接相邻地图块后的边缘配接



## D. 运行一个点查区数据库

3. 108. 通过构建一个地理数据库，国家统计局办公室开始认识到围绕一个地理模型来构建自己的好处。为了设计人口普查数据模型，国家统计局办公室必须考虑其数据用户的需要，以及法令和宪法对人口普查工作的规定。一般说来，可以开发一个概念模型，以便连接收集的基本地理特征，这些特征来自点查层，例如点查区、物理特征(包括海拔、斜率和形貌)以及影响人口普查工作的其他层，包括点和面积特征，例如地标。上面提到了向量拓扑结构发挥的重要作用，尤其是在边界和行政区的划分中。

3. 109. 虽然向量拓扑结构很重要，其他数据格式可能也需要被引入一个地理信息系统项目，包括光栅文件格式的扫描地图、卫星图像、航摄照片、CAD文件(其引入时可以带拓扑结构和属性或者像没有特征的“意大利面条”一样)以及以一串串点表示的全球定位系统数据。

3. 110. 1990年代末和2000年代初计算机技术的进步之一是在计算机技术领域开发了基于对象的数据。在面向对象的编程中，对象可以被定义为有差不多几千个不同的特点，被输入一个地理信息系统时，这些对象的“举止”在模拟环境下是可以预知的，可以建立复杂的模型，例如旅行路线分析。

3. 111. 地理数据库不仅仅是电子数据表。实体类型可以被定义为具有特殊的特性，这些特性支配着在现实世界中的行为。点查区，作为一个地理单位，是一种对象，其功能是为开展人口普查-民意调查活动对领土进行划界。在形态学上，点查区是连续的，嵌在行政单位中，由以人口为基础的单位组成。

### 1. 关系型数据库

3. 112. 在讨论人口普查地理信息系统数据库的具体结构前，应首先回顾一下关系型数据库的概念。所有大型的可操作的地理信息系统都建立在地理数据库的基础上，正如可提出证据加以证明的一样，地理数据库是地理信息系统最重要的部分。地理数据库构成所有查询、分析和决策的基础。一个数据库管理系统(DBMS)就是储存数据库的地方。

3. 113. 一个地理对象可以被定义为几何、特性和方法综合在一起的一个完整的東西。同一综合类别的对象被分为不同的对象类别，每一类中的单个对象被称为“实例”。在许多地理信息系统中，每个对象类别被存储为一张数据表，每一行一个对象，每个特征一列。例如，对象类别“街灯”可以包括“瓦斯灯”、“钠气灯”和“汞气灯”等实例。

3. 114. 关系型数据库模型用以存储、检索和处理列有坐标数据库地理特征的数据表，其依据是实体关系模型。

3. 115. 在地理环境内，“实体”可以是行政管理或人口普查单位，也可以是汇编各种特点的其他任何空间特征。例如，实体可以代表“点查区”的特

征(图III.17)。某个区或国家的单个点查区即是这种实体的实例，并且将在该实体表的行中表示。相反，实体类别指数据库表的结构：存储在表列内的实体的属性。对于某个点查区来说，实体类别可以是赋予点查区的独特标识符、地表面、人口、小组负责人区域的代码等。应当注意的是，实体类别仅指数据库表的一般定义，不是指为每一实例记录的实际值。实体类别方面的一种或多种属性(列)用作关键字或标识符。其中的一个是主关键字，用作实体类别的独特标识符。对于某个点查区数据库而言，这个关键字是点查区的代码。

图III.17

#### 实体表实例——点查区

实体：点查区  
类别(属性)

点查区代码	地区	人口	小组负责人代码
723101	32.1	763	88
723102	28.4	593	88
723103	19.1	838	88
723201	34.6	832	88
723202	25.7	632	89
723203	28.3	839	89
723204	12.4	388	89
...	...	...	...

实例

主关键字

#### 各种关系型数据库和地理数据库结构

3.116. 数据库管理系统(DBMS)可以分为各种类型，包括关系型、对象型和对象-关系型。关系型数据库管理系统(RDBMS)包含一系列表格，每张表格都是一个二维阵列，记录着研究对象的属性。虽然RDBMS灵活而有用，但设计它们不是为了处理丰富的数据类型，例如地理，其中的对象拓扑结构和关系可能很复杂。商业的开放源RDBMS程序包括Microsoft Access和Oracle。

3.117. 对象型数据库管理系统(ODBMS)的设计是为了解决RDBMS的一个重要缺陷，即它们不能直接在数据库中存储全部对象。ODBMS可以持久稳固地存储对象，并提供面向对象的查询工具。对象-关系型数据库管理系统(ORDBMS)是对象型数据库和关系型数据库的混合，包含一个关系型数据库引擎以及一个用于处理对象的可扩展性框架。在理想的情况下，一个ORDBMS由以下部分组成：用于SQL(结构化查询语言)查询的查询剖析器、查询优化器、查询语言、索引、存储管理、交易服务和复制。

3.118. 通过使用地理DBMS扩展，软件公司回应了对它们的关系型数据库中空间功能的需求。这些是带空间数据库扩展的大DBMS。例子包括IBM DB 2 Spatial Extender、Informix Spatial Datablade和Oracle Spatial。这些软件产品处理点、线和多边形，把它们作为可以纳入更丰富的类型的特征类型，利用的是拓扑结构和线性参照功能、R-树索引和四叉树方法。

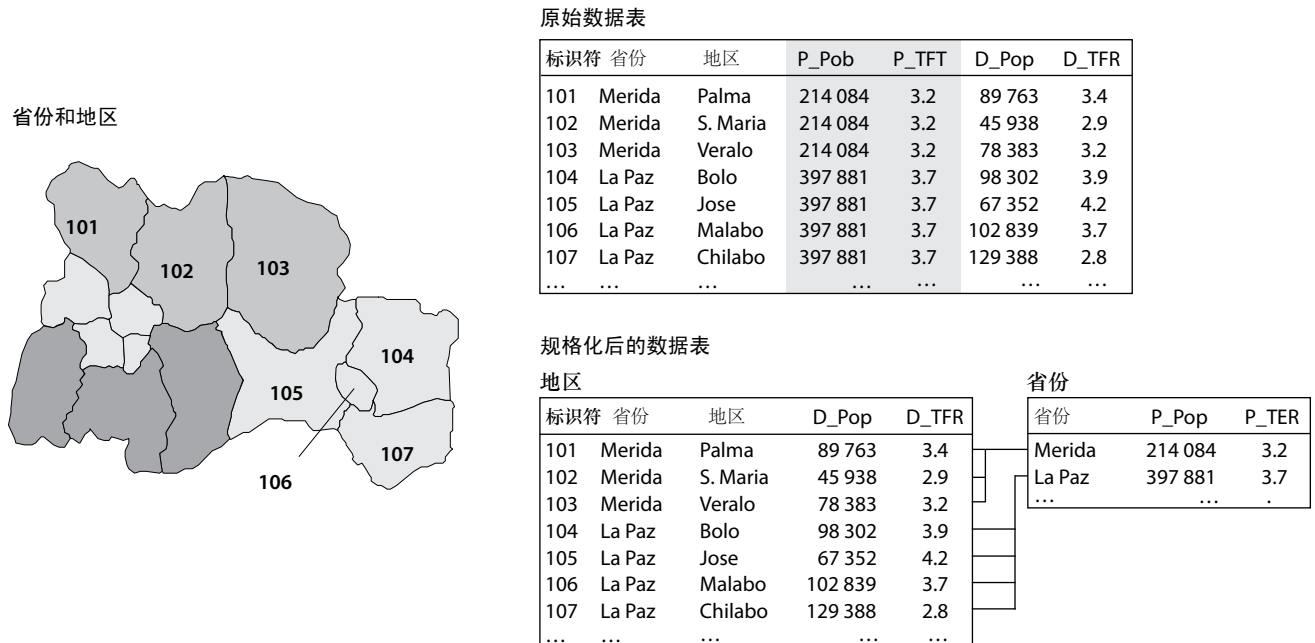
3. 119. 在对DBMS进行扩展以存储地理数据中存在特殊的挑战。对象-数据模型是以几何为中心的，因为它们是把世界作为一堆对象加以塑造的，例如点、线、多边形或光栅。作为单独的程序，利用编程或脚本在几何图形上进行操作。不过对于地理系统来说，这过于简单了，尤其是在地理系统包含许多实体，而这些实体有大量特性、复杂的关系和行为的时候。

3. 120. 各种关系确定了各实体间的关联性。例如，某个描述点查区的表格可链接实体小组负责人区域的表格。该表列入了各种属性，如组长姓名、地区主管机构及联络资料。该表中的主关键字是组长代码(CL代码)，这一代码也列入点查区表格内。这样，关系型数据库管理系统可以结合两个表格，以便点查区表格内的每种实例符合小组负责人区域表格内的实例。

3. 121. 通过一系列步骤设计关系型数据库结构的过程叫做“规格化”。规格化的结果是将数据库的冗余数据减少到最低程度。换句话说，数据被列入许多表格中，因此避免了多次重复出现的数值。这样便减少了存储空间，避免了在插入、删除或更新等标准数据库操作过程中可能发生的误差。

3. 122. 图III. 18利用地区数据库的例子说明简单数据表及其规格化形式之间的区别。在第一种情况下，省份信息重复地用于省内各地区。这不仅浪费了存储空间，而且使得更新或改变省份信息更加难以进行。每个地区的数值都必须予以替换。在规格化数据库结构中，省份的名称用压缩的数字代码代替，这样便能够连接另一个表。在这里，省份代码是省份信息的主关键字，这些信

图III.18  
关系型数据表



息包括省份名称、人口及总生育率。利用省份代码临时合并两个数据表后，即可存取省份信息，用于地区表中的每种情况。

3. 123. 界定明确的数据库结构并非是无紧要的工作。有些数据库管理方案提供了能够自动创建关系型数据库结构的规格化功能。不过，这通常不能很好地替代总数据库的综合设计。

## 2. 数据库内容(数据建模)的定义

3. 124. 一旦确定人口普查地理活动的范围，人口普查办公室必须非常详细地界定并记录地理数据库的结构。这个过程有时被称作数据建模，涉及列入数据库的地理特征的定义、这些地理特征的属性及其与其他特征的关系。最终的产出是详尽的数据字典，不仅可以指导数据库开发过程，以后也可作为文献资料使用。

3. 125. 应当指出，许多地理信息系统数据库是在没有详尽的数据建模的情况下创建的。这个步骤需要时间和某种程度的数据库概念专门技术。综合人口普查制图项目追加投资是有理由的。数据建模过程必须严格和一致，因为做到这一点才能保证数据库达到高质量并且易于维护。对于首次完成这个过程的人口普查制图机构来说，聘用经验丰富的地理信息系统数据库顾问全程指导工作小组的工作是可取的。

3. 126. 前文已讨论过，许多国家和国际机构一直在积极开发空间信息类属数据模型，作为国家空间数据基础结构的一个组成部分。人口普查办公室通常能够轻易地让国家空间数据基础结构的标准适用于统计数据收集的具体需求。如果收集不到这类信息，数据模型的开发必须在机构内部进行。其他国家的绘图或统计机构制作的模板可为数据模型的开发提供有用的参考。

3. 127. 附件三举例说明了数据字典可能对数据模型做出的解释。与数据模型有关的内容是下节将要讨论的元数据标准，以及同出售给公众的数据库并行使用的简化数据库字典(见附件四)。

## E. 数据质量问题

### 1. 准确度要求

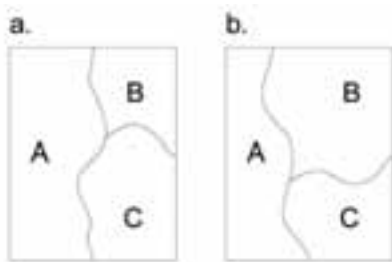
3. 128. 在规划数字数据库开发项目过程中，制定可接受的数据准确度标准可能是最重要的工作之一。在公用事业和设施管理、地形或水文绘图等诸多领域，已经制定了数据库的数据准确度标准，可供任何新项目采用。相反，人口普查制图通常采用手工技术和草图在短时间内进行，几乎不关心地理准确度的问题。如果人口普查地图仅用于普查目的，这种做法是足够的。但是，采用地理信息系统技术后，人口普查地图成为政府、私营及学术部门的许多分析应用程序的必要组成部分。所以说首先投资进行数字人口普查制图有正当理由，这是一个主要因素。当人口普查地图与其他数字地理数据源结合使用时，

准确度方面存在的缺点立即表现出来。因此，数字人口普查制图的准确度要求高于传统的人口普查制图技术。

3. 129. 地理信息系统的准确度指的是属性数据(地理属性表及其列入的人口普查数据)和地理数据。与属性数据准确度有关的问题与人口普查数据输入及其处理工作中碰到的问题相同。因此，本文仅作简要讨论。地理数据准确度与存储在地理信息系统数据库中的对地球表面特征进行描述的点、线和面积有关。

3. 130. 地理数据准确度可分为“逻辑”和“定位”准确度。定位准确度有时也称为绝对准确度。逻辑准确度指地理特征间的完整关系。例如，某个地理信息系统数据库层的道路须连接另一个层面的桥梁。存储在水文数据库并且界定两个行政单位间的边界的河流应与这两个单位间的边界吻合。在地理信息系统数据库内用点表示的某个城镇应位于另一地理信息系统层的对应行政单位内。同样的逻辑关系可在外观截然不同的地图上正确表述。例如，在图III. 19上，两份地图正确地表述了三个行政单位之间的相邻关系。

图III.19  
逻辑准确度



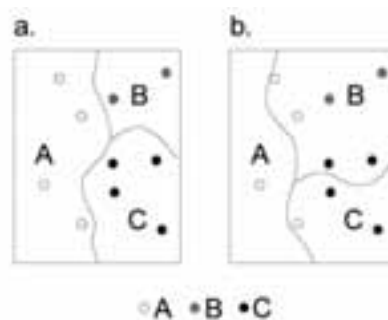
3. 131. 相反，定位准确度要求地理信息系统数据库内的特征坐标与它们在地球表面的真正位置完全对应。这就是说，采用全球定位系统等精确测量工具进行绘图测量的精度必须非常高。当然，不发生定位错误的数据集也能够精确地表述地理特征间的逻辑关系。

3. 132. 在有些应用中，逻辑准确度比定位准确度更重要。对于某个人口普查数据库而言，可能更为重要的是了解某条街道可界定一个点查区的边界，而不是了解精确的坐标可以非常精确地表述真实世界的道路方位。事实上，在传统人口普查绘图工作期间制作的草图一般具有逻辑准确度，但定位准确度较差。如果地图仅用于保障人口普查的点查，只要图像失真不会使点查区的定向工作无法进行，那就不成问题。但是，如果人口普查地图随后用作其他用途，将产生严重问题。

3. 133. 例如，图III. 20a利用准确度非常高的全球定位系统给出了一组已经确定的抽样调查点。下方的底图定位准确度非常高。因此，点的位置在正确的行政单位内。相反，图III. 20b的底图虽说具有逻辑准确度，但定位准确



图III.20  
不保持定位准确度将产生的问题



度差。因此，有些准确测量的全球定位系统点位落在了错误的行政单位内。其后果是，如果按行政单位汇总调查反馈，结果将不正确。

3. 134. 因此，如果得到的边界的用途超出了实际点查范围，数字化人口普查绘图过程的目标应当是达到十分高的定位准确度。当然，极少数地理数据收集的准确度能够达到100%。在所有制图工作中，无论是手工制图还是数字化制图，必须对能够达到的准确度与数据质量达到这一程度所需的时间和资金进行权衡。一般来说，将准确度提高到90%或95%以上时，投入的时间及其他资源的比例则远远高出准确度的提高比例。事实上，有人估算，准确度从95%提高到100%所需的资金占项目全部预算额的95% (Hohl 1998年)。

3. 135. 依据点位的位置界定准确度标准是地形绘图工作常用的方法。例如，标高点的高度离实际方位的距离应小于 $x$ 米，在所有情况下误差应小于 $y\%$ 。允许误差随着地图比例尺的减小而增加。例如，比例尺为1:25 000的地图，其误差应小于比例尺为1:100 000的地图。由于人口普查地图在很大程度上以现有的地形图为依据，因此应与国家绘图机构的专家密切合作，制定人口普查绘图准确度标准。这种做法可保证人口普查绘图项目的产品质量符合其他国家数据地图系列的产品质量。

3. 136. 尽管定位准确度要求较高，但准确度标准过于苛刻带来的后果则是成本的提高、用户的期望值过大以及制图人员可能因无法实现过高的目标而灰心丧气。准确度标准过低的后果是产品质量达不到要求。如果用户了解产品的局限性，要么就会拒绝使用产品，要么会过分信任地使用，后一种做法可能产生的后果是分析结果发生严重误差。开发地理信息系统数据库时采用的普遍概念是“适用性”。这种观念考虑到了数字空间数据库永远无法达到完美这一事实。虽然数字空间数据库可能会适用于某项工作，但若用于另外一项工作，其质量也许达不到要求。

3. 137. 在确定质量标准时，人口普查单位不仅应考虑内部需求，而且应考虑数字化人口普查地图的外部用户的需求。因此，应与所有利益攸关方协作制定数据准确度准则，作为用户需求评估的一个组成部分。可对标准产生影响的因素还有可用资源、原始材料的质量——不同数据层的质量不同——以及进行实地数据收集选用的技术。

## 2. 质量控制

3. 138. 质量控制是保证在人口普查制图过程中开发的数据库符合规定的准确度标准的过程和规范。修订后的《人口与住房普查原则和建议》(联合国, 2008年)强调质量控制的重要性, 并且概述了人口普查过程中存在的质量控制问题。这些普遍概念也适用于构建地理数据库。

3. 139. 质量控制必须贯穿整个人口普查过程, 对地理方案也同样重要。作为一项基本战略, 最好的方案设计包括以确保人口普查工作各个阶段的质量控制为目的的协议。

3. 140. 检验和误差校验是质量控制过程的关键。不过, 质量控制也关系到人口普查制图人员在数据转换过程的各个步骤控制误差的态度问题。应该鼓励人口普查人员报告产品输出发生的问题。问题反复发生可能说明程序不完备或者培训不达标, 并且可能要求变换制图人员的任务、改装设备或换用其他技术。因此, 重要的是, 制图人员要敢于报告自己工作中出现的问题, 明确质量控制程序的总体目标。

3. 141. 虽然制图人员从事不同的专项工作在大多数情况下能够提高数据的总体质量, 但开发地理信息系统数据库的许多工作重复性非常强。从事某项单调工作的制图人员会因为注意力分散而增加误差, 轮换工作任务可防止这种现象的发生。这样做能够让制图人员接触整个数据转换过程的各个方面, 因此会提高他们对工作和产品总体质量的了解程度。另外, 应要求制图人员建议改用有利于提高数据质量的规程。应当在受控环境下而不是日常工作过程中评估这些更改建议, 然后方可付诸实施。因此, 尽可能使数据质量达到最好是一个长期过程。

3. 142. 质量控制程序包括自动和人工方法。由于自动化程序迅捷、可靠, 因此最好选用。不过, 数据转换的许多方面只能通过用眼睛检查和比较来评估。自动地理属性数据技术类似于人口普查数据输入采用的技术。范围和代码检查可保证属性字段仅包含允许使用的数值。可以进行统计测试以确定异常值。数字数据库内的行政或人口普查单位数目须与地理面积主目录内的对应数目一致。地理面积标识符是地理数据库最重要的单字段, 因为它保证了数字底图和汇总后的人口普查数据相一致。因此, 在属性数据检查——无论是自动的还是手工的——方面投入的最多资源应致力于保证这一属性没有误差。

3. 143. 地理数据自动质量控制选择方法相对有限。有些地理信息系统软件包将检查数据库的拓扑准确度, 例如, 各地区是否封闭, 各条线之间是否相连。村庄数据库可列入质量已知的行政单位边界数据集, 以保证村庄数据库的行政标识符正确无误(多边形内的点运算)。有些误差很明显, 例如两个单独数字化的行政单位的边界不吻合。其他误差不大容易发现, 例如, 有些内边界或道路不在地理信息系统数据集中。因此, 地图产品的质量控制在极大程度上须依赖对原始材料(地图、航摄照片等)和数字化数据的目视比较。为此, 印制数字地图采用的比例尺最好与原始地图一样。然后, 既可将原始资料与印制的地

图放在一起进行比较,也可将它们叠放在看板台上进行比较。任何有规律的误差均说明数据转换程序有问题,应当立即加以解决。切勿由编制数据的制图人员进行手工误差检查。

3.144. 应编制记录质量控制步骤的资料。一般来说,硬拷贝日志表是记录数据质量的最恰当的手段,尽管也可以采用自动化的数字表格。日志表应载明已执行的质量控制程序、执行时间、执行人、经核对的数据的编制人以及检验结果。应编制手工及自动检验日志。这些日志不仅记录数据的准确度及其系属特征,而且能够指出哪些制图人员须继续接受培训。

3.145. 连贯一致的质量控制程序应该能使最终产品的准确度令人满意。但大多数项目通常还要实施一个最终的质量保证步骤,这个步骤包括进行另一轮检查和解决问题的最后过程。有关质量保证的问题将在下文讨论。

### 3. 将国家领土划分为操作单元

3.146. 完整的数字点查区数据库由数以千计的单元组成。对于国土面积较大的国家,将所有点查区多边形存储在同一物理数据层内通常不切合实际。相反,国家领土可分成多个操作区。这样,在权力下放的人口普查行政结构内,不同地区办公室及同一地区办公室的不同操作人员可以开展合作,同时创建数据库的各个部分。如果国家数据库的各子数据库的边界能够取得一致,随后即可合并各子数据库,制作区、省或国家地图。不过,由于这个过程涉及手工连接跨越两个或两个以上的片区的连接特征,因此要求进行一些边缘配接工作。

3.147. 对于国土面积较大的国家,制图工作权力有可能下放。在这种情况下,操作区自然由各地区人口普查办公室负责普查的区域界定。例如,某个国家可将人口普查制图工作交给四个区级办公室,总办公室同时履行总体协调机构的职能,并作为其中的一个区级机构。在每个区级普查办公室内,数据库可进一步分为小型区域。小型数据库的工作通常对计算技术的需求不大。将数据库分为许多小的部分还能够让若干操作人员同时操作数据库的不同部分。

### 4. 数字化的行政底图

3.148. 操作设计方面的选择取决于处理情况以及组织环境。如果决定采取权力下放的方法,国家人口普查办公室首先应创建适用于国家主要行政级别的国家边界模板。例如,人口普查办公室应创建、购买或委托制作省、区数字空间边界,最好包括分区边界。这些边界的准确度应当非常高,并且应配有使它们适用于点查区大比例尺(例如,比例尺至少为1:250 000)制图的详细说明。这些边界应用于整个人口普查制图过程,并且应用于向这些行政级别发布空间基准汇总人口普查信息。

3.149. 国家制图机构可能已经采用数字方式绘制出这类边界。如果是这样,这些边界将代表官方认可的国家数字行政底图(见前文关于NSDI的讨论)。显然,行政底图使用的代码应与人口普查数据库使用的代码一致。

3. 150. 各操作区的正式边界应发布给主管点查区划界工作的机构。然后将点查区边界输入这些正式行政单位多边形内。这样做将可保证在随后进行任何汇总时，相邻区的边界完全吻合。如果区边界分别由各地方机构数字化，不可能完全吻合，因此要进一步进行大量编辑工作。此外，重复工作量很大，原因是同一条边界须分别由相邻地区普查办公室或操作人员进行两次数字化处理。

## 5. 处理不相交的地区单元

3. 151. 行政单位常常分为独特的单独空间单元或多边形。例如，某个地区可能由陆地部分和众多岛屿组成。对于人口普查数据处理而言，这种现象不是问题，因为适用于该地区的每个人口普查数据表仅有一份记录。但在地理属性数据库中，该地区将有两份或多份记录——每个多边形一份记录。如果人口普查属性信息通过地理属性表连接各个多边形，这种现象将会产生一些问题。在关系型数据库系统中，人口普查数据记录连接地理信息系统数据库中地区标识符相同的每个多边形。绘图平均值或密度不构成问题。整个地区的平均收入或人口密度相同。不过，如果某个用户想汇总各地区的人口总数，人口总数或住户数目等计数数据将构成问题。由于记录重复用于属于同一地区的每个多边形，有时会发生重复计数问题，最后得出的总数将过大。解决这个问题的方法有两种。

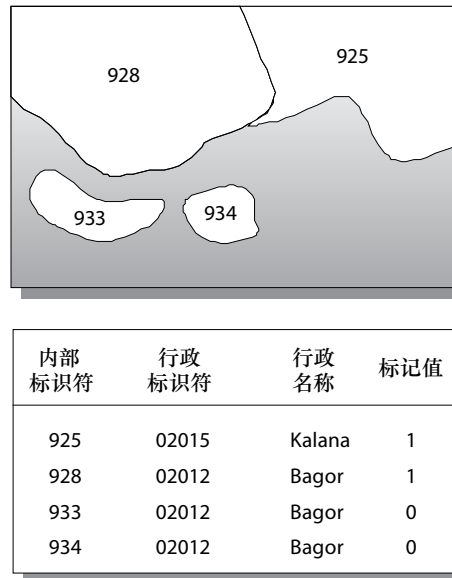
3. 152. 有些先进的地理信息系统软件包可给“地区”下定义。地区可由一个或多个单个多边形组成，但在地理属性表中，每个地区仅有一份记录。系统可在自身内部跟踪哪个单个多边形属于哪个地区。在有些软件包中，各地区甚至可以重叠，尽管这不是人口普查应用程序的有用特征，因为点查区必须互不重叠。

3. 153. 许多低端地理信息系统软件没有这个选项。在这种情况下，简单的解决方案是向地理属性表添加其他数据字段(标记值)(见图III. 21)。这个字段假定属于该地区的最大多边形的值为1，较小的多边形的值为0。在汇总或平均任何属性值之前，用户首先可以仅选择在这个字段中的值为1的多边形。另外可以添加包括属于同一单元的多边形的数目的字段。使用地理信息系统软件包的频率功能或交叉列表功能可以迅速生成这一信息。

## 6. 面积计算

3. 154. 如果人口普查数据库列入一些标准地理变量，其效用将得以提高。其中最重要的变量是各点查区或行政单位的面积。如果数据库能够正确地以相同的面积基准投影为基准，所有地理信息系统软件包都将计算多边形的面积。不过，依数字化边界的分辨率和准确度而定，地理信息系统测量结果的误差可能很大，原因是边界的模糊程度非常高，以及有些岛屿过小而无法列入小比例尺地图。因此，如果有条件的话，最好使用国家制图机构编制的准确度较高的面积数据。

图III.21  
处理由若干多边形组成的行政单位



3. 155. 面积数据用来编制密度估算值，最重要的是编制人口密度估算值。已公布的面积数据通常指行政单位的所有法定边界内的面积，即总面积。有时，面积数据可在某种程度上致使密度估算值令人产生误解。举例来说，国家人口普查出版物报告了邻近一个大湖边的几个地区的面积。报告的面积包括湖岸到湖心的那部分面积(见图III. 22)。某些地区的总面积列入湖泊面积后翻了一番。因此，面积翻一番这一因素致使实际人口密度的估算值过低。如果官方人口密度统计数据用作，例如，资源分配的标准或用来确定政府计划的可行性，人口密度的定义可产生重大影响。

3. 156. 在存在上述问题的国家，人口普查办公室可决定报告两个面积区域：一个是行政单位的总面积，一个是土地面积。土地面积指总面积减去水面面积及其他无人居住区，如自然保护区的面积。一些国家还报告农业用地的面

图III.22  
占若干行政单位一大部分面积的湖泊



积。它可让用户计算农业人口密度，反之亦然，也可以计算出该地区居民人均占有耕地的公顷数目。利用恰当的地理数据层可以非常轻松地在地信息系统内计算出这些面积数据，但须做出与上文提及的制图泛化有关的解释，防止产生误解。无论在哪种情况下，编制详尽的文件记录净面积的定义都是重要的。

3. 157. 由于大多数地理信息系统软件包将数据库的每个多边形看作一项单独记录，为由一个以上的多边形组成的行政或人口普查单元计算的地理信息系统面积数据对密度计算没有用处。相反，同属一个行政或人口普查单元的所有多边形的面积须予以汇总。利用适当的交叉列表功能在地信息系统内可做到这一点。

## F. 元数据的开发

3. 158. 本《手册》认为，地理数据库的构建应被看作是一个长期的过程，而不是一次性的工作。在长时间内，数据库的各元素将被反复存取，有时要经历一段较长的过渡期。工作人员可能频繁变更，这就意味着机构存储器的依据必须比参与初始数据开发的地理空间分析家的记忆力可靠。因此，务必编制文件，详细记录下开发数字空间人口普查数据库的各个步骤。

3. 159. 说明数据质量、格式和处理步骤的信息以及其他与数据集有关的信息称为元数据，或者“关于数据的数据”。元数据的作用如下：

- 支持对某个单位拥有的数字数据集的维护和更新。
- 通过提供资料说明某个数据集是否适用于外部用户，支持数据的销售。
- 支持外部编制的数据集与单位数据的集成。

3. 160. 哪些是必不可少的元数据，不同的编制人员有不同的看法。因此，许多国家着手制定通用地理元数据标准，目的是统一记录空间信息的规范。因此，这些国家通过促进空间数据交换和集成来支持国家空间数据基础结构的开发。各国都有团体在制定空间元数据标准，一些国际性组织试图协调这项工作。其中的一些组织有：国际标准组织化地理信息/测绘工作小组 ([www.statkart.no/isotc211/](http://www.statkart.no/isotc211/))、欧洲委员会公开信息交换服务中心 ([www2.echo.lu/oii/en/oii-home.html](http://www2.echo.lu/oii/en/oii-home.html))以及亚洲和太平洋地区地理信息系统基础结构常设委员会 ([www.permcom.apgis.gov.au](http://www.permcom.apgis.gov.au))。

3. 161. 由于空间基准人口普查数据是国家空间数据基础结构不可或缺的组成部分，数字人口普查地图的制作应尽可能与国家的其他地理空间工作相结合。关于元数据，如果国家和地区已出台元数据标准的话，就意味着国家人口普查机构应该采用这些标准。与国家主管当局——通常是国家制图机构或部门间咨询委员会——开展密切合作将促进这些标准的推行。如果国家标准没有出台，人口普查单位可采用其他国家的适用标准节省时间和资源，而不是从头开始制定元数据标准。

3.162. 一个元数据标准的例子是，美国联邦地理数据委员会(www.fidgc.gov)制定的数字地理空间元数据内容标准。这个标准用以说明元数据数据库载有的信息的类型。全套标准具有综合性，各专门委员会负责制定具体数据类型的准则。例如，文化和人口普查数据小组委员会就设在美国人口普查局(<http://www.census.gov/geo/www/standards/scdd/>)(另见美国联邦地理数据委员会1997年b)。这里仅讨论元数据定义的主要组成部分。

3.163. 数字地理空间元数据内容标准由七个主要部分组成，可被看作一种数据库模板，通过模板的字段描述空间数据集的不同方面。有些字段载有事先界定的一组代码或属性。但许多元素是文本字段，数据编制人员用它们说明数据库的特征，如质量或系属信息。最重要的元素具有强制性，必不可少，因此，应将它们输入每个数据集。这种强制性字段集对于定义人口普查单位的元数据模板来说是一个良好的开端。其他字段则标有“如果适用的话，具有强制性”或“供选用”的说明。

3.164. 标准的主要组成部分有：

- 标识信息，包括数据集的名称、覆盖面积、关键字、用途、摘要以及存取和使用限制。
- 数据质量信息，如水平和垂直准确度评估、逻辑一致性、语义准确度、时间信息、数据集的完整性以及系属。系属包括用于编制数据集的数据源以及处理步骤和中间产品。
- 空间数据组织信息，指存储点、光栅、向量和数字地图平铺信息等数据的方式。
- 空间基准信息，包括地图投影及界定坐标系的一切相关参数。
- 实体和属性信息，载有数据集属性的详细定义，包括属性数据类型、允许值和定义，基本上同上文3.127段所述数据字典载有的信息一样。
- 分配信息，包括数据分配器、数据文档格式、脱机媒体类型、在线连接数据、费用及订购过程。
- 元数据参考信息，提供有关元数据本身的信息，最重要的信息是谁创建了元数据及创建时间。

3.165. 除了这七个主要部分外，信息标准还包括三个次要元素。这些元素经常用做主要部分的基准。仅需将这些元素存储在某个位置，不用多次重复。这三个次要部分是：

- 引文信息，确保创作者、标题、出版日期及出版者作为基准的一致性。
- 时间信息，包括某一日期、多个日期或日期范围。
- 联络信息，如联络人和/或组织、地址、电话和电子邮件。

3.166. 对政府和其他数据编制机构间的元数据信息进行规范化的一个优势是，可以开发管理和使用元数据的类属系统。例如，管理数字地理空间元

数据内容标准的工具很多。其中包括文本输入形式、数据库或网页浏览器格式(通过因特网或局域网)以及可供图书馆或因特网数据分配系统使用的元数据阅读器。商业软件销售商也向他们的软件中添加文件编制工具,便于采用数字地理空间元数据内容标准格式编制元数据。

3.167. 用于人口普查绘图项目的元数据的模板定义仅仅是元数据管理的一个方面。另一个方面是元数据维护程序的执行。人口普查组织须确定输入数据的时间和人员、数据存储格式(纸型或数字文档)以及监督所得信息是否具有完整性、准确度及可用性的人员。元数据编制应与创建数据库的各个步骤同时进行,不应仅把它看成是文件编制的最后一个步骤。为了维护数据的未来用户或外部用户的利益,元数据应被视作同空间数据库本身一样重要。

## G. 概要和结论

3.168. 第三章提供了关于组合一个数字化点查区级别的数据库的各个步骤的技术内容,包括地理数据库的基础、数据输入技巧、地理编码和点查区划界。

3.169. 到这一操作阶段结束时,国家统计局办公室在上一次人口普查的基础上,创建了一个无缝的全面覆盖的点查区,准备进行实地更新。在某个时刻,国家统计局办公室已经完成了自己在办公室所能做的一切,现在必须走出办公室,实地开展工作。

3.170. 最后的任务是支持为人口普查-民意调查创建一个点查区级别的数据库的决定。这将使得人口普查更为准确,并为后来的分析和发布提供好得多的输入数据。这对人口普查后的结果发布也会起到支持作用,尤其是对人道主义方面的用途而言(例如灾害防范和管理),这个问题将在第六章谈到。





## 第 四 章

### 利用全球定位系统和 遥感数据整合实地工作

4.1. 本章继续探讨在第三章中介绍过的构建点查区级别的地理数据库的各个步骤。在这里，通过直接谈及新的技术和数据来源，我们认识到因为卫星技术——即全球定位系统和遥感(包括航摄照片)——而成为可能的新的工具和数据来源的价值。

4.2. 第四章的主要话题是利用全球定位系统和遥感来进行点查区划界工作。这样做的主要理由是实地验证在国家统计办公室的地理信息系统实验室根据以前的人口普查地图确定的点查区边界。不过，在没有正确度的地图可用时，在为完成和验证开展实地工作之前，要在人口普查的主要办公室制作正确的地图，作为点查区划界工作的基础。借助于遥感数据，地理分析员可以查明最需要更新的地域，并将这些地区与需要最小限度的更新的地区区分开来。将介绍一些全球定位系统要素，以及关于在人口普查中使用全球定位系统数据(包括将全球定位系统用于点查区和行政边界的划分以及住宅单位和集体生活区的定位)和提取其他特征的一些准则。将谈及移动计算和便携式计算机。将从卫星图像和航摄照片方面论述遥感技术，为国家统计办公室使用该技术提出一些要素和准则。

4.3. 在理想的情况下，工作进行到这一步，统计机构已经扫描了以前的人口普查绘制的点查区地图，并且已经把这些地图转化为一个数字化地理点查区数据库。不过，尽管花费了大量时间，并且似乎很全面，但该数据库事实上是一个粗略的草稿，因为在人口普查主要办公室绘制的点查区地图还没有经过实地更新。总部与人口普查实地办公室在这项工作中合作的程序将取决于人口普查活动的集中化程度，以及通信结构和国家的可到达性。本《手册》认为总部和实地的活动将通过以实地为基础的活动以及数据共享结合在一起。

4.4. 总的说来，地理数字化最重要的目标是利用新技术来更快地绘制出更好的地图并提高人口普查数据的整体质量。同样，我们把重点放在国家的特殊条件上，关注如何将领土划分为各个行政区和点查区，供统计机构在人口普查中使用。通过综合卫星图像，分析人员和人口普查规划者可以确定需要补充进行实地工作的地区，例如考虑到城市周边地区有了新的增长。出于规划和后勤的目的，有必要提早查明这些重点地区，以便确定自上次人口普查以来发生快速变化的地区，并把重点放在这些地区。所谓的“变化情况探查”方法就是

这个意思，利用实地、实验室和遥感数据综合的办法，可以非常有效地完成这一工作。

4.5. 遥感的定义是利用图像传感器技术来收集关于特定地区或对象的信息。遥感是“看”人口普查地形的一个有力手段，可以极大地提高点查的准确度，对国家统计办公室来说，这也可能是一项巨大投资。遥感所能做的不只是增加实地地图绘制；实际上，最好将遥感的使用与其他数据来源相结合，例如实地地图、边界描述和以前人口普查的点查区。由于遥感是一笔巨大的投资，国家统计办公室在着手获取图像和进行培训之前需要审慎制定计划。这样一个计划将预先规定数据分析和产品，这样国家统计办公室就不会制成对公众来说没有多大用处的无关数据。当技术被看作是昂贵的附加物，而不是一种达到目的的手段时，这一点特别重要。一旦国家统计办公室回答了数据的特殊用途的问题，才能够从资源、尤其是人力资源的角度对影响进行评估。

## A. 全球定位系统 (GPS)

4.6. 曾经作为一件舶来品的全球定位系统使导航领域发生了翻天覆地的变化，如今已是随处可见。近年来，全球定位系统也使人口普查实地绘图领域发生了一场革命。通过以一种易于使用的格式记录经纬度，全球定位系统使得将位置添加进任何应用软件成为一件轻而易举的事。由于全球定位系统接收器价格的下降——不到100美元就可以买到一个可靠的型号——全球定位系统被用于许多领域，并且有了大量个人用户，大举进入了汽车、船舶、建筑和农场设备消费者市场，并被安装在手提电脑和膝上型电脑中。最大的专业用户群在公用事业管理、电信、勘察和导航等领域。但是，全球定位系统还促使生物、林业、地质等领域的实地研究取得了进展，并且还不断应用于流行病及人口研究领域。此外，全球定位系统日益成为人口普查绘图应用领域的重要工具。

4.7. 大多数讨论提到的所谓的“美国系统”指的就是全球定位系统。全球定位系统是使用最为广泛的系统，它的使用为接收器制造商和勘察机构培育了一个大型商业市场。关于其他卫星定位系统，包括俄罗斯的全球导航卫星 (GLONASS) 系统、欧洲联盟的伽利略 (Galileo) 系统以及中国的北斗 (Beidou) 系统，下文也将进行简单的介绍和讨论。

### 1. 全球定位系统的工作方式

4.8. 全球定位系统是利用轨道卫星的一个全世界的无线电导航系统。全球定位系统接收器收集由24颗卫星——21颗有源卫星和3颗备用卫星——及其地面站发出的信号。名为NAVSTAR的系统由美国国防部进行维护。卫星在高度约为20 000公里的六个轨道平面上围绕地球旋转。在任何特定时间都有5至8颗全球定位系统卫星位于某个地面全球定位系统接收器的“视野范围”。

4.9. 全球定位系统接收器在地面上的位置通过测量地面到几颗卫星的距离确定，用三维(X、Y和Z)方式表示。全球定位系统卫星和接收器分别发出

一个精确的同步信号(利用所谓的伪随机码)。能够实现同步的原因在于卫星和接收器的时钟非常精确。接收器能够测出内部信号和从卫星接收的信号之间的延迟时间。这种延迟是信号从卫星传送到接收器的所需时间。由于信号以光速(299 338公里/秒)传输, 若想知道距离, 只要用延迟时间乘以光速即可。

4. 10. 如果知道了若干颗卫星的距离, 方位即可通过三角测量方法确定。如果我们测出另一颗卫星的距离, 则可将我们所处方位的范围缩小到两个圆周的交点。为了证实我们所处的确切位置, 我们应确定到第三颗卫星的距离。三颗卫星周围的距离圆周仅有一个交会点, 这一点就是我们所处的确切位置。当然, 在现实生活中, 我们生活在三维世界中。如果仅能测定到一颗卫星的距离, 我们所处的方位可能在卫星圆周的任何地方。如果测定了两个距离, 我们可能位于两个圆周相交后形成的圈内。最后, 第三颗卫星的圆周与这个圈的交会点有两个。当然, 这两点当中通常仅有一点切合实际。不过, 为了提高方位估算值, 因此进行第四次测量。第四次测量也有助于纠正接收器内部时钟目前存在的走时不精确的现象。相反, 卫星上的原子时钟走时非常精确。

## 2. 全球定位系统的准确度

4. 11. 全球定位系统接收器的价格不贵, 可提供比较准确的信息, 显示出用户在地球上随时随地的纬度、经度和高度。据大多数低成本“自主的”全球定位系统接收器销售商声称, 在民用方面, 所记录的方位通常能精确到15-20米之间。高度信息在某种程度上不如经纬度可靠。

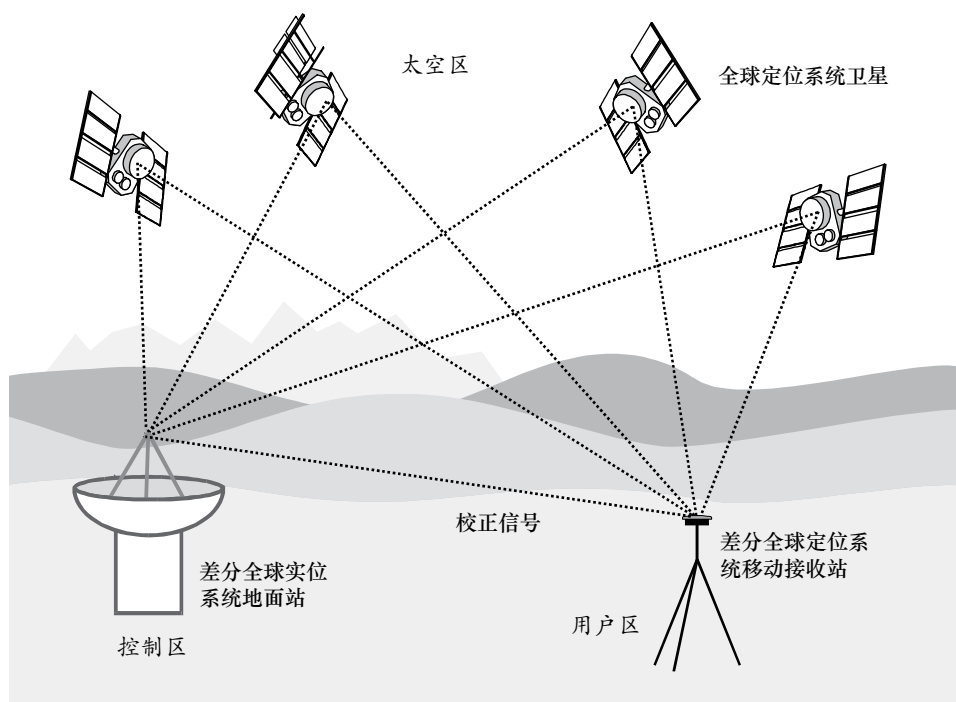
4. 12. 影响准确度的因素有几个。其中的一个因素是卫星的数目和方位。在理想的情况下, 为了使几何计算极为准确, 卫星应遍布天空。定位精度因子指的是天空中卫星的分布, 可以用一个数目进行量化。另一个影响信号质量的因素是轨道不规则。可以通过卫星时间表来说明不规则情况。另一个因素是月亮的影响, 如改变穿过大气层的信号的大气干扰。“多路误差”是由建筑物或其他固体物体散射信号引起的。这种误差或多或少地表现为随机噪声(即方位上的随机和短期变动)。最大的误差源是选择可用性, 这是美国国防部为降低信号准确度而制造的, 在2000年以后逐步停止了这种做法。尽管选择可用性的做法被中止了, 但在战争时期可能会恢复。

4. 13. 全球定位系统坐标的重复读数不一定能提高坐标估算值。为了获得更准确的方位, 必须在长时间(24小时以上)内平均坐标读数。实际上, 有更好的选择方法能够改进全球定位系统坐标。

## 3. 差分全球定位系统

4. 14. 对于准确度要求较高的应用而言, 差分全球定位系统使用基站传输的校正信息, 并采用已知的精确坐标纠正卫星信号(见图IV. 1)。差分全球定位系统正在变得日益普及。差分全球定位系统需要两个接收器之间的合作, 一个是固定的, 一个是漫游的。差分全球定位系统基站和移动全球定位系统设备收到的信号会产生同样的误差。基站接收器测量计时误差, 然后向移动全球定

图IV.1  
差分全球定位系统



位系统设施发送正确信息。基站接收相同的全球定位系统信号，然后计算全球定位系统信号的旅行时间并同实际情况进行对比，计算出误差校正系数。利用差分全球定位系统可以达到的准确度取决于系统和坐标收集程序。利用非常廉价的硬件在较短的观察时间内能够达到的准确度为2米，在静止的情况下结果甚至更好。利用较昂贵的系统并延长收集每个坐标读数数据的时间能够使准确度达到亚米。

4. 15. 实施全球定位系统实时校正的选择方法很多。许多国家的政府机构目前正在安装差分全球定位系统基站，不断广播校正信息。这种基站一般位于海岸区，用于海上导航目的。费用相对低廉的差分全球定位系统基站有时由各类用户群建设，例如说精细耕作农业领域的用户群。此外，有些便携式高端全球定位系统装置可被改装成广播校正信息的差分全球定位系统基站，这种基站越来越多地利用互联网。用户需要在附近找到一个确切知道的位置，然后就可以对其绘制精确的地图。最后，校正信息也可由对地静止卫星广播。例如用于飞机导航目的。

4. 16. 如果实时数据不必要，那么全球定位系统坐标的后处理也许是一种有用和不那么复杂的选择。在这种情况下，用户使用标准全球定位系统接收器收集坐标。对每个坐标来说，接收器存储器记录了时间和使用的卫星。回到办公室后，用户可下载当时的校正信息，并将校正因数应用于收集到的所有坐标。在许多国家，校正数据文件的来源是一些商业和公共机构。如果不能从二级渠道获取这类信息，可在中央位置建造一座差分全球定位系统基站。例如，为了支持人口普查绘图工作的进行，可在首府建造一座差分全球定位系统基

站，以便采用费用低廉的标准接收器在实地收集的坐标数据能够在随后得到校正。在较大的国家，将不得不建立多个基站。

4. 17. 一些值得一提的新的国家系统利用新的卫星技术和互联网上的通信。广域增强系统或WAAS是美国联邦航空局建立的一个大陆性的差分全球定位系统。它包括一个对地静止(这个术语指的是卫星的位置是固定的，与轨道运行卫星正好相反)卫星，该卫星利用美国境内运行的24个站点，广播关于一个全球定位系统频率的校正信息。持续运行的基站或(CORS)也是为航空目的设计的，但还有许多其他用途。每个站点都提供全球定位系统测量值，这样就可以进行误差校正，并且使得定位准确度接近几个厘米。在出版本《手册》时这一服务只在北美提供，尽管在编写这一版本的时候很快就可以向非洲推出一个类似的系统。

4. 18. 其他地区的政府正在开发类似的以卫星为基础的差分系统。在亚洲，日本正在开发一个多功能卫星增强系统或MSAS，该系统将在全亚洲运行，准确度为3米。欧洲有EGNOS，或欧洲对地静止导航重叠服务，包括三个对地静止卫星和一个由34个地面站组成的网络。EGNOS开始于2005年，将在2008年得到准确度小于2米的认定。服务区域包括非洲和南美洲。

#### 4. 其他全球卫星导航系统

4. 19. 除了美国的NAVSTAR全球定位系统外还有几个其他选择。这类系统的一个更笼统的术语是全球导航系统或GNS。俄罗斯的全球定位系统是全球导航卫星系统(GLONASS)，由俄罗斯联邦国防部操作。始于1976年的GLONASS到1995年完成了整个星群，但随后失修，2007年只有7颗卫星在轨道上。俄罗斯空间局计划在2011年以前将GLONASS恢复到全面部署状态，即24颗卫星。与印度政府签署了一项合作协定，将从印度领土上发射两颗卫星，作为回报，印度可以接收到高精确度的信号。

4. 20. 欧洲联盟的伽利略系统正由欧盟和欧洲空间局打造，计划在2013年以前全面投入运行。伽利略的特色是将在2006年至2010年发射的30颗卫星组成的星群，位于慕尼黑和罗马的两个地面站。与美国运行的全球定位系统相比，所做的一点改进将是包括“完整性信息”，直接告知用户信号误差；另一点是伽利略将在极地范围运行。伽利略将与美国的全球定位系统在用户层面实现共同使用。

4. 21. 中国提出的北斗系统最终将包括35颗卫星，包括5颗对地静止卫星和30颗轨道卫星。将向中国用户提供准确度为10米的免费服务；外部订户付费可以得到更准确的服务。2007年将再发射两颗北斗II卫星。

4. 22. 商业上可获得的全球定位系统接收器在价格和功能方面各不相同。技术规格决定可以达到的定位准确度。接收器的功能越强大，价格就越高。用户需要权衡准确度的提高与费用的增加是否物有所值。在许多绘图应用中，标准系统的准确度就足够了。接收器在下述方面也各不相同：方便用户、

导航中有用的跟踪功能——许多接收器如今可以绘制地图——以及支持地图投影和地理基准系统。在选择全球定位系统接收器中还要考虑耐用性、耗电情况（因为电池价格昂贵，用于汽车的AC适配器可以发挥作用）、坐标存储能力以及将存储的坐标转入一台膝上型电脑或桌上型电脑容易度。

4. 23. 大多数销售商提供将全球定位系统接收器与一个手提或笔记本电脑相结合的集成产品，这样就可以直接在屏幕上画出捕获的坐标，或者孤立地显示，或者显示在一张数字化底图上。下面将介绍这些技术以及集成的实地制图系统。

## 5. 人口普查制图应用领域使用的全球定位系统

4. 24. 全球定位系统技术在制图活动中有许多用途，包括为人口普查活动编制和校正点查员地图。如上所述，我们必须强调将新技术的使用纳入一个可靠和详细的全面计划的重要性。有了差分全球定位系统，便可以确定点查区边界的精确地理方位，并且可以具有成本效益的方式获得诸如服务机构或村庄中心等点状特征的方位。坐标可以下载，或用手工方式输入地理信息系统或其他数字化制图系统，并可在这些系统中与其他地理基准信息合并。下面列出了实地划分点查区边界的准则，还举出了可以利用全球定位系统的具体人口普查活动的例子（关于点查区划界的更多信息见第三章）。

4. 25. 正如我们所知道的，对于人口普查数据的收集来说，点查区是与操作有关的地理单位。点查区也可作为人口普查数据的发布单位使用，但它们主要用于收集数据。点查区的特征包括它们完全覆盖国家领土。点查区的用意是代表“相同人口规模”的区域。它们使用道路和水体等特征作为边界，这些特征在地面是可以观察到的。除了来自以前人口普查的地图外，利用遥感数据将节省大量工作时间。实地检查可以保持在最低水平，资源可用于处理快速变化的地区。

4. 26. 在点查区划界工作中，人口普查代表必须明白理想的点查区规模包括地区规模和其中的人口规模。一个点查区代表一个点查员在人口普查数据收集期间能够覆盖的地域大小。一个点查区的人口限值以国家统计局制定的计划为基础，利用一次人口普查预备调查的结果来确定点查所需的天数。

4. 27. 人口估计数代表点查区划界的最关键因素。点查区不但以面积而且以人口为基础，因此，需要确定用于估计每个点查区内的人口数的方法，尽量提高工作的准确度，并与资源相一致。如果不能提前准确估计点查区人口数，那么就会牵制点查工作，威胁结果的质量。

4. 28. 通过与地方官员合作可以估计点查区的人口数。尽管可能对点查区单位本身不熟悉，但农村地区的地方官员可以估计小村庄和村庄的大小。有时，就小地域而言，对住宅单位的估计比对人口的估计容易。如果没有地方官员的参与，国家统计局工作人员在实地调查中或者通过利用现有信息，例如航摄照片、卫星图像、公路图或规划图、行政记录、人口登记簿、公用事业

公司记录或以前的人口普查结果，可以做出估计。在最后一种情况下，也许需要对数字进行调整，以反映该地区的人口增长。

4. 29. 在全面覆盖这一最重要的目标基础上制定了点查区划界准则，这意味着重叠和缺口要尽量减少。通常，一个特殊符号，例如一条弯弯曲曲的线起到强调的作用。点查区边界跟随醒目的特征，例如公路、溪流、湖泊和铁道线，不管是在地图上还是在地面上，这些特征都很明显。一些地貌特征，如山脊线和森林，不是理想的边界，应该在没有别的地貌特征的情况下使用。“分支线”可用来表明需要将公路两旁的住宅单位包括进来。

4. 30. 总的说来，在点查区划界中，应权衡人口规模与土地面积以及旅行难易程度。如果旅行条件需要点查员花费过多时间从一个住宅单位到另一个住宅单位，那么点查区的规模应缩减。自然特征，例如悬崖、河流、沼泽和林地可能会成为旅行的障碍，人为的条件也一样，例如散居模式、四分五裂的公路系统以及基础设施的总体情况不佳。

## 6. 一些与全球定位系统有关的特殊的制图任务

4. 31. 一些与全球定位系统有关的特殊的地图绘制任务包括：

- (a) **点查区边界的划分。**点查区边界是多边形，并且以自然的地貌特征为基础。鉴于点查区的规模小，利用全球定位系统来划分全部土地而不采用分类法，可能不太实际，分类法就是把工作的重点放在近来发生变化的地区。在遥感科学中，这被称为“变化情况探查”法。如果一个有2 000万人口的国家需要40 000个点查区，每个点查区500人，那么用全球定位系统装置记录所有边界的工作可能需要花费许多年的实践。最好的办法是根据以前的人口普查对点查区边界进行数字化，并且只在需要的地方利用基于地面的全球定位系统方案，尤其是在边界发生变化时，例如设立新区或者土地合并；
- (b) **行政边界的划分。**在大多数国家，行政边界(例如省、区和分区)已经以小规模(即一般化)的方式划分好了。很可能这些单位缺乏人口普查工作所必需的精确度。国家统计办公室应仔细权衡在人口普查时进行仔细的行政边界的划分可能带来的好处与时间和劳动方面的成本。如果有可能，国家统计办公室应该获取关于行政边界的现有数字化文件。有可能从国家制图机构获取这样的文件。应查询随附的元数据，包括资料和投影信息，然后才可以在地理信息系统项目中利用这些数据；
- (c) **住宅单位的位置。**一些国家甚至对国内的每一个住宅单位记录了经纬度，有时甚至对每一处住房拍摄了照片。在全国范围内开展这样的工作需要耗费大量资源。如果人口普查工作者在点查期间从事这



项工作，可能不需要花费大量时间，但必须完整地储存这些文件并编制索引，以避免代价高昂的重复劳动；

- (d) **集体生活区的位置。**集体生活区是一种公共住宅或机构住宅，包括旅馆、军营、孤儿院、工地宿舍、男修道院、女修道院、老人院、医院、宿舍和刑事机构。住在公共机构里的人有时最容易遭受自然灾害的危害，人道主义规划者越来越多地要求了解公共生活区的地理位置，以便有效规划灾害应对措施。此外，用全球定位系统装置确定集体生活区的位置可能比测量国内所有住宅单位轻松，因为它们的数目较少；
- (e) **其他相关特征(包括公路)。**诸如公路之类的特征可能对划分点查区边界或提供导航信息有用。水体也对人口普查工作者确定方向有用，地标可以用作对卫星图像或附属地图确定地理基准的控制点。国家统计局办公室应请求政府提供基础数据的数字版，例如公路，这样可以为实际点查节省时间和金钱。

4.32. 就用于人口普查而言，应仔细考虑广泛使用地理信息系统的问题。对许多任务来说，包括记录每一个住宅单位和点查区边界的读数，大量实地工作人员所需的设备可能超出了一个人口普查项目的资源限制。

4.33. 全球定位系统坐标用于人口普查制图的确切方式因选用的人口普查制图策略不同而异。全球定位系统可以点状模式收集坐标，例如，某个村庄的每栋建筑物或某个城市街道网络的交叉路口的坐标。在数据收集过程中使用的地图或绘制的草图有助于回到办公室后判读坐标信息。另一种可能性是采用流模式收集全球定位系统坐标。在这种情况下，系统定期记录坐标。这样，步行、开车或骑自行车沿着道路行进便可自动记录线状特征。如果计划周密，这种做法对于创建街道或道路网络数据库来说可能具有成本效益，尽管它取决于选用的数据质量标准，即所得线条的准确度是否足够。出于安全考虑，也是为了在没有正常电力供应的地区提供备份，从全球定位系统装置中检索坐标，并以手工方式记录在数据表中，这是一种成本较低的选择。如果电池没电了，全球定位系统装置就不能运行，因此，应确保充满电，包括几块后备电池。应锁定全球定位系统装置的其他功能，这样实地工作人员就不能对它们重新编制程序或者改变其设置。最后，应该给装置做明显标记，因此，如果被盗，这些装置不容易被转卖。

## 7. 对使用全球定位系统的培训要求

4.34. 一个全球定位系统项目要想获得成功，国家统计局办公室必须审慎管理设备的采购，安排培训和人员需要以及制定点收集协议(Montana和Spencer, 2004年)。建议为此类活动指定一名全球定位系统协调员。至少，这样一个人应了解全球定位系统装置及与收集和储存点有关的信息。协调员还应监督实地工作者，确保其工作的一致性。针对实地工作者的培训方案可能包括理解全球定位系统的运行，以及接收器如何计算位置以及装置的故障检修。

4. 35. 在理想的情况下，应提前几年，作为人口普查总体规划的一部分，规划全球定位系统的使用。至少，应在实施前六个月完成与全球定位系统有关的项目计划，并且制定和通告数据收集策略。在规划过程的这一点上，可以确定硬件方面的需要。可以设计数据表，包括关于不同单位和地理特征的一致的命名协议，以便在数据库中点查区或特征的标识符与全球定位系统的路点编码相一致。

4. 36. 在全球定位系统硬件或备份严重缺失的情况下，全球定位系统协调员应执行一个备份制度。这可以简单到在调查表格纸张上写下经纬度，然后抄入一个电子数据表。

4. 37. 为了将全球定位系统路点载入一台膝上型电脑或桌上型电脑，可以利用GPS Utility、EasyGPS等共享软件和免费软件GPSBabel，与制造商的软件一起来上载点。操作人员可以在Excel或其他电子数据表中重新确定上载的文件格式，以便纳入一个地理信息系统项目，在这个点上可以添加补充的属性数据。在一些程序中，全球定位系统可以以经过正确格式化的表格或“事件数据”的形式被添加到一个地理信息系统项目中。

## 8. 总结：全球定位系统的优缺点

4. 38. 全球定位系统的优点如下：

- 实地数据收集的部件价格相对低廉、易于使用。现代化的装置几乎不要求进行培训就可投入使用。
- 对许多人口普查制图用途来说准确度足够了——采用差分校正方法，可以达到高准确度。
- 收集到的数据可直接读入地理信息系统数据库，不必要进行中间数据输入或数据转换步骤。
- 全球通用。
- 未来五年内新的全球定位系统将变成在线的。

4. 39. 全球定位系统的缺点如下：

- 全球定位系统装置虽然廉价，但如果规划不当，利用全球定位系统开展的实地工作可能非常费时(因此也费钱)。
- 全面规划包括确定广泛使用全球定位系统将得到什么产品。
- 在密度大的城区或林区，信号可能受阻(多路径误差)。
- 标准全球定位系统准确度也许不能满足城市地区及捕获线性特征的要求，必须采用差分技术。
- 差分全球定位系统价格昂贵，在许多偏远地方可能没有，另外，实地数据收集所需时间较长，提高信息准确度的后处理较为复杂。
- 短期内的数据收集也得投入大量全球定位系统设备，使得广泛使用全球定位系统可能非常昂贵。

- 全球定位系统装置越复杂，需要的培训就越多。

4. 40. 在全球定位系统的应用中，可能会发生各种各样的问题。在密度大的城区，多路径误差可能使得几乎不能界定邻近的点查区。由于卫星信号无法穿过固体物体，高楼林立或街道两边林木繁茂使得难以接收大量卫星发出的信号。通过步行到较为开阔的地带并弥补所记录的坐标的偏差，受过培训的数据收集员仍然能够获得坐标信息。在某些情况下，必须使用差分全球定位系统或用其他数据源，如已出版的地图、航摄照片，甚至用在实地工作期间绘制的草图核对全球定位系统读数。有些国家开发了全球定位系统基站系统，利用差分全球定位系统支持高准确度绘图工作。在一些发展中国家，这样的网络尚不存在。

## B. 利用便携式计算机的集成的实地制图系统

4. 41. 在一些国家，将个人电脑的功能与全球定位系统相结合的新技术被用于人口普查活动。利用手提电脑或个人数码助理(PDA)的一个好处是“直接捕获”——能够直接记录信息，而不需抄写，省却了几个中间步骤。地图可立即得到更新。坐标一经捕获，即刻便显示在便携式计算机的屏幕上。如果计算机内存有数字底图，坐标可显示在数字底图上方。实地工作人员可以添加任何需要的属性信息，并将这些数据存储在地理数据库中。然后回到办公室后可将该信息输入地理数据库。由于笔记本电脑及其他手提电脑越来越廉价，集成的实地绘图系统正在成为为人口普查进行的实地数据收集的可行选择。

4. 42. 技术的进步，包括全球定位系统、无线通信和计算机小型化，使得用于便携式地理信息系统的新应用软件能够大量出现，尤其是专门用于人口普查实地工作的软件的开发。大量袖珍装置支持许多显示、查询和简单分析应用软件，程序和数据储存在内存里，因为便携式装置没有硬盘驱动器。其他相关特征包括通过蓝牙和/或WiFi无线连通性通信，借助个人电脑实现同步化，以快速上传数据和更新数据，并确保备份，以防数据丢失。为在“智能手机”上使用开发了地理信息系统。加固型PDA可用于极端环境下的移动数据收集。用于便携式电脑的软件包括流行办公软件的“简陋”版。用于手提电脑的地理信息系统软件的例子包括Autodesk OnSite、ESRI ArcPad和Intergraph Intelliwhere。

4. 43. 打算将手提电脑用于人口普查-制图前或实际点查工作的国家统计局办公室必须考虑成本问题。装备全球定位系统的手提电脑的价格大都在750美元以上——对于加固型电脑来说价格更高。对那些管理地理活动的人来说，实施方面的问题是：比起一个100美元的全球定位系统装置来，使用一个750美元(或者更贵)的装置有什么好处？一些因素包括屏幕上地图可读性、能量需求，尤其是在没有可靠供电的地区，以及对于电脑、甚至加固型电脑不利的其他环境条件。

## 方框IV.1

## 全球定位系统方面的经验案例研究：斐济

斐济2007年的人口普查是太平洋地区首次使用全球定位系统技术来连接人口普查问卷与国内所有住户确定了地理基准的位置。在斐济，正如在许多发展中国家，没有准确的住处位置数据库，并且缺乏街道名称和编号意味着没有可用的地址清单。全球定位系统有一个明显的优势，因为它提高了住户定位的准确度和覆盖面，并且被用作管理和监测工具(用于检查及数据确认)，因为对于人口普查管理者和点查员来说可以精简人口普查活动。全球定位系统的定位使得人口普查数据的加总可以被纳入一系列管理单位，例如健康、教育和环境管理单位。

为开展人口普查活动大约采购了200个Garmin eTrex全球定位系统装置和20台膝上型电脑；对10名官员进行了全球定位系统和ESRI Arcview方面的培训，他们负责进行有效性检查和下载路点；在三周内培训了200名全球定位系统操作员。全球定位系统操作员和主管人得到了一张简表，上面逐步概述了如何安装全球定位系统装置，记录路点，从全球定位系统上下载路点到计算机上，在图像上显示路点以及输出和打印JPEG图像文件。

开展了实地工作，这样全球定位系统操作员就能跟随点查员并收集路点。主管人定期到这些操作员那里检查并下载数据。如果可以得到图像，在Google Earth中，路点被置于图像之上，以验证位置。设计了一个“三次记录路点数”的方法，把住户位置与调查问卷相连接。点查员首先把一个带唯一的六位数编码的不干胶贴纸放在接受了调查的房屋上，如果房屋有围墙，则在大门上也放一张。一张小的带相同号码的不干胶贴纸被放在关于那家人的表格上。如果前门和大门相隔很远的话，则把一张备用的不干胶贴纸放在住户的大门上。这有利于获取路点。其次，全球定位系统操作员访问住户，并输入该数目，作为路点编号，以便随后将其与调查问卷相连接。

一旦完成了实地工作，将为路点创建两个数据库，一个涉及其经纬度坐标，另一个涉及调查问卷。可以利用惟一路点/调查问卷编号连接两个数据库。在这个阶段要进行检查，以确定调查问卷缺失和不正确的路点。

遇到了一些困难，可以从中吸取教训，以利以后的人口普查工作：

- 在全球定位系统点收集的早期，操作员在达到可接受的精确度水平之前就在获取路点了。为了解决这个问题，增加了等候时间。
- 两个宗教节日导致了不干胶贴纸的遗失(为庆祝节日重新粉刷了房屋)，这意味着操作员必须回到办公室收集居民的姓名和地址，然后再回到实地收集路点。
- 作为路点输入全球定位系统和从调查问卷输入数据库的六位数编码有时不正确。在实地对此进行补救的一个办法是未来使用附着在全球定位系统接收器上的条形码阅读器。
- 另一个错误的缘由是全球定位系统路点采集工作落后于点查工作，有时落后数月。如果点查小组也采集路点，那么这些错误就可以消除。

对斐济来说，这种全球定位系统信息一个预期的好处是能够防范和管理自然灾害。将基本的住户信息覆盖在一个数字化地形模型上非常清楚地显示出受到任何特定灾害影响的人口。

资料来源：在Nouméa举行的讲习班上斐济所做的介绍，2008年。

图IV.2  
在全色卫星图像上划分点查区边界



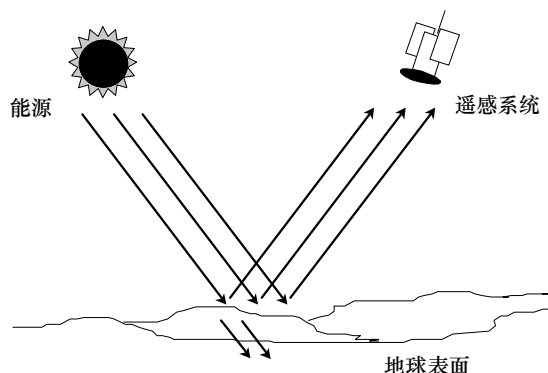
### C. 卫星遥感

#### 1. 利用图像来实地核查在人口普查总部制作的点查区地图

4.44. 自2000年的《手册》出版以来，遥感数据已经变得更多、更普遍、更容易使用。尤其是自高(一米或更高)空间分辨率的卫星图像出现以来，遥感技术使地图绘制领域发生了翻天覆地的变化。利用这一宝贵资源来开展人口普查工作的机会已经成熟。对国家统计局来说，一个挑战是要绘制地图的领土幅员广阔。如果实际加以使用，卫星图像能够节省大量人员和时间，因为国家统计局只需把重点放在关键地区。遥感数据可用于独立检查实地核查程序。本《手册》提倡的一种方法是对一个国家的表面积进行筛选，把它分为需要较多和较少关注的地区。这就是所谓的“变化情况探查”法，对于确定居民区的周界特别有用。

4.45. 根据第三章中介绍的方法，在这个阶段，国家统计局已经对旧的人口普查点查区地图进行了数字化，覆盖了其他地理数据，并全部输入了地理数据库，但还没有实地修正结果。随着临时的点查区边界覆盖在遥感图像上(见图IV.2)，可以很快确定人口定居点的位置以及优先地区。不过，对此类活动的规划必须详细和务实。

图IV.3  
遥感过程

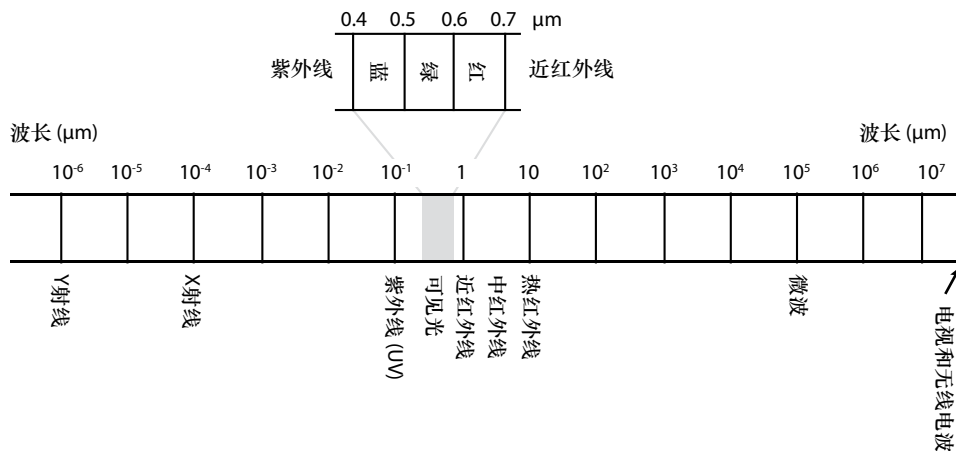


## 2. 卫星遥感的原理

4. 46. 在地理信息科学中，卫星图像是众多的地理数据形式之一，可以用于分析和显示结果。图像的格式被安排为一个连续的域，也称作光栅，其中数据行对应代表数值的像素。遥感被认为是主要的数据捕获形式；事实上，它是最流行的主要光栅捕获形式。自轨道运行卫星出现以来，在提取地球表面的读数方面有了巨大的潜力，有许多现成的和规划的应用软件可用于人口普查工作。对于人口普查工作来说，遥感技术的主要优势之一是，如果操作正确，它能覆盖危险的或无法达到的地区，从而节省大量实地工作的时间。

4. 47. 卫星图像从基于空间的系统获得，其中的大多数图像通过被动光学传感器收集而来，这种传感器对地球表面物体的可见和不可见的电磁波谱反射进行测定(见图IV. 3和IV. 4)。与主动的传感器，例如雷达(雷达也能穿透云层)相比，大多数卫星数据的收集被认为是被动的，因为它接收发射自地球的能量。卫星系统不用摄影胶片记录反射的能量，而采用一种类似电荷耦合器件摄像机的光电检测阵元件对电磁辐射强度进行测量，并将其作为一个数字化规则光栅或行列图像记录下来。

图IV.4  
电磁波谱



4. 48. 卫星传感器是按多谱或全色模式进行工作的。多谱的意思是卫星收集几种图像(或频带), 其中的每一个图像仅对电磁波谱中某一部分(通常是可见光和近红外线范围内的部分)的反射能量进行测量。把一幅图像分解成不同的频谱带和将不同的频谱带合成为一幅图像的分析能力有助于依据地面上不同地理特征的反射特性对其进行分类。例如, 稻田会在某一特定的带宽内显示出强烈信号, 而建筑群则会在另一带宽上更清晰地得到表现。全色卫星传感器能够在很宽的频谱范围内捕捉反射能量。所得的图像类似于黑白摄影照片。这种图像的分辨率通常高于多谱图像, 因此更适合作为人口普查绘图的基础。

4. 49. 传感器系统得到的数字化数据包括数字矩阵, 表示地球表面相应位置上的反射能量级别。卫星将这些数据发送给地面接收站系统, 并在系统内对数据进行辐射测量校正和几何校正以及地理基准处理。处理后的数字化图像或打印图像可以通过观看进行解释, 类似于航摄照片解读, 或者可以利用地理空间技术对其进行分析或者将其与地理信息系统项目内的其他层合并。卫星数字图像可以显示在地理信息系统上, 技术熟练的操作人员可以对图像上的特征进行界定。对于许多用途来说, 例如观测土地利用情况或自然资源管理, 则需使用统计技术对多谱图像进行分类。这些土地覆盖区预测分类是基于已知种类的控制现场与其频谱特性的校准关系。

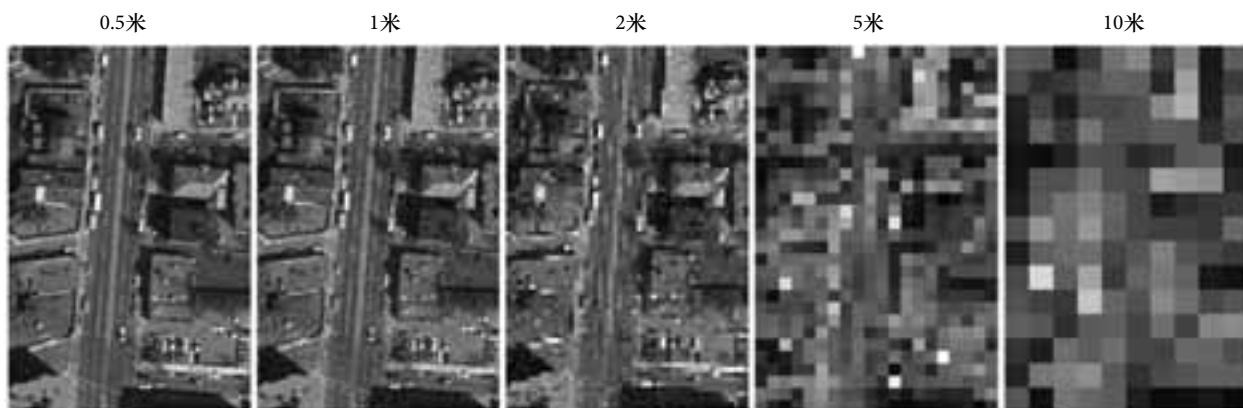
### 3. 遥感数据的分辨率

4. 50. 卫星图像的空间分辨率是按地面上一个像素的大小测定的。还有其他的分辨率度量标准, 包括时间的、辐射测量的和光谱的。就商用卫星来说, 像素大小不等, 最流行的高分辨率系统, 如Quickbird和Ikonos, 其分辨率小于1米。印度遥感(Indian Remote Sensing)、地球观测实验系统(Spot)的全色传感器和大地卫星的多谱图像也被认为是高分辨率系统, 可以绘制比例为1:25 000到1:50 000或者更小的地图。2008年3月, 遥感公司GeoEye计划发射一个像素大小为0.41米的传感器。

4. 51. 图IV. 5比较了各种像素尺寸, 它们是从一幅数字航摄照片上模拟下来的, 该图片的分辨率为0.5米, 其他尺寸则是通过加总形成。图像所覆盖

图IV.5

航摄照片和卫星图像像素大小图解



的地面面积为100x150米。在分辨率为2米的情况下，每一座房屋，甚至每一辆汽车都可以辨别出来，不过像素尺寸如果再大就区分不出来了。从遥感数据中还可以抽取更多的信息，需要使用先进的图像处理技术，包括边缘检测算法和专用过滤算法。此类技术已经成功地运用于地图绘制，并且用于探测发展中国家一些快速发展的城市中新开发区发生的变化。

4. 52. 在可以利用的各种图像中，从低空间分辨率/高时间分辨率的METEOSAT和雷达到中等空间和时间分辨率AVHRR图像，卫星图像可用于多种目的。服务于人口普查工作的遥感产品相对较少，即高空间分辨率数据，例如Quickbird(0.82米)、Ikonos(1米)、IRS(5.8米pan)、Orbimage 3和4(1米)，以及SPOT 5(2.5米)。Ikonos是1999年发射的，Quickbird是2001年发射的。

4. 53. 对于大多数人口普查应用来说，5米或更高的空间分辨率对于确定住宅单位和人口定居点的范围是必需的，多谱图像则不是绝对必要的。这些数据源的不利之处是高分辨率图像的覆盖面较小，这意味着需要购买大量图像来覆盖即使一国领土的一部分，使得完全覆盖的前景十分昂贵。

4. 54. Enhanced Thematic Mapper (ETM) 的空间分辨率为30米或15米，因此，Landsat可以识别线性特征，例如公路和河流，以及其他补充层，例如湖泊和其他水体。Landsat还有一个很大的优势，那就是它是免费的。在某些情况下，Landsat和ASTER可能对记录土地覆盖层的变化和人类的影响有用；这些数据源对于人口普查工作来说不那么有用。在决定使用哪种遥感产品进行点查工作时，国家统计局办公室应仔细评估目标和成本。一个可以考虑的选择是，尤其对大国而言，使用组合的办法，对人口密集区，比如城市，采用1米或5米的图像，用Landsat覆盖周围的农村地区。在这种情况下，农村地区仍需进行实地的地图绘制，以便更好地规划点查期间的后勤工作。

4. 55. 大多数商业经营者对获取卫星图像提供几种选择。一般说来，定价的基础是图像是“归档的”还是首次收集的。最昂贵的选择是专门要求对某一个地区紧急购买图像。分辨率较高的卫星覆盖的地面面积较小，因此它们只能在飞行路线上覆盖选定的地区。一种不那么昂贵的选择是利用图像档案不那么紧急地获取图像，可以以低得多的价格购买一部分图像。图像的价格还取决于对原始数据的处理程度。这可能包括辐射测量校正、几何校正和带或者不带地面控制点的地理基准处理。例如，可以从Digital Globe得到的归档的图像(2008年3月)的价格为每平方公里(km<sup>2</sup>)16美元，最少要订购25平方公里。Ikonos图像一般比较便宜，约为每平方公里8美元。价格将取决于购买的图像所覆盖的面积大小，面积越大，每平方公里的价格就越便宜。与根据卫星图像制作的数字化正射照片地图相比，原始图像数据要便宜很多。不过，一般说来，购买的图像是经过充分处理的。表IV.1提供了空间分辨率非常高的民用卫星遥感产品的清单。联合国秘书处外层空间事务办公室保存有一张更完整的清单(见www.oosa.unvienna.org)。其他传感器包括ALOS(这是在美洲使用的日本系统)、Alaska Satellite Facility(ASF)和CBS26, 以及IMPE(对非洲免费的中国-巴西系统)。



表IV.1  
空间分辨率非常高的民用卫星遥感产品

产 品	公 司	发 射	模 式	最低点的 像素大小	高度 (公里)
- iQ1	IKONOS	2001	4m/4m	0.61/2.44	450
IKONOS 2	IKONOS	1999	4m/4m	0.82/3.28	680
- VI3	AVIRIS	2003	4m/4m	1.0/4.0	470
S15	SPOT	2002	4m/4m	5.2/5.10	830
Ca. Q1	NASA/日本	2004	4m	2.5	617
Ca. Q2	NASA/日本	2004/5	4m	1	630

4. 56. 可以在线或通过一个地方或区域经销商订购图像。另一个也许更便宜的办法是利用国家的空间数据基础结构来进入公共图像档案。国家统计局办公室的代表可以联系其他机构，特别是国家制图机构，以索取现有遥感档案，用于人口普查工作，必要时可签署一个使用协议，明确规定如何使用图像。

#### 4. 卫星遥感数据在线资源

4. 57. 正如可提出证据加以证明的那样，对于国家统计局办公室来说，新的以因特网为基础的卫星图像资源同样重要，这些图像可作为人口普查工作的直观辅助物，同时免去了购买图像的费用和麻烦。遥感数据如今可以在线获取，利用的应用软件如Google Earth、ArcGIS Explorer、Microsoft Virtual Earth和其他在线数据资源。与购买图像相比，利用在线数据检索的好处是国家统计办公室可以测试图像的适用性，而不必进行前期投资。不利之处是分辨率和图像的总质量可能不适合绘制详细的点查区地图。

4. 58. Google Earth是一个虚拟地球程序，通过编制卫星图像的目录并加以显示来绘制地球。Google Earth对地理空间领域产生了影响，并且正在引起公众对卫星技术和地图的兴趣。该应用软件由Keyhole开发，2004年被Google购得。包括在Google Earth中的大多数图像来自Digital Globe的Quickbird，尽管一些航摄图像和三维建筑如今也包括在里面。Google Earth对于世界大部分地区提供15米或更高的空间分辨率，利用的是地理投影和WGS84资料。Google Earth上的图像受版权保护。用户不能访问真正的数据来源，只能查看。不过，他们可以添加自己的数据。

4. 59. 目前，Google Earth提供三个层次的授权：免费的Google Earth阅读器，一个Plus版(订购价为每年20美元)以及一个商用的Pro版(订购价为每年400美元)(2008年3月的价格；核实最新定价)。Plus版包括全球定位系统集成，用户可以阅读来自全球定位系统装置的轨迹和路点。另外，Google Earth Plus提供对Magellan和Garmin全球定位系统产品系列的直接支持、较高的分辨率打印、通过电子邮件支持客户以及能够从一张电子数据表中读取地址点的数据输入器，该数据输入器使用的是逗号分割数值，但只限于100个点/地址。

4. 60. Pro版本的功能包括附加软件, 例如电影制作, 允许用户利用三维绘图工具展示以位置为基础的数据。用户还可以从一张电子数据表中迁移最多2 500个点位置。一个地理信息系统数据输入模块让用户可以以图形文件和.tab格式添加地理数据, 包括人口数据。Pro版可用的测量工具允许用户计算面积和直线的距离, 并且可以输出11x17英寸或4 800像素的图像。Google Earth Pro不在线销售, 必须通过销售代理购买。

4. 61. 在任何版本的Google Earth中, .kml(Keyhole标记语言)格式的地理数据可以输入到Google Earth中。一个可免费下载脚本可将点和多边形转换为适当的格式, 尽管多边形可能更难输入。初步的拓扑结构, 即没有多边形属性表或节点位置, 意味着有问题的数据不得输入。Pro版本可以输入但不能输出图形文件。用于人口普查的扫描得到的地图, 例如点查区地图, 如果输入Google Earth的话, 可能存在问题, 因为对文件大小有相当严格的限制(18 000x18 000像素)。通过使用一个“regionator” Python脚本可以克服这一点, 但可能在文件管理方面带来问题。

4. 62. Google Earth, 尤其是带图像输出功能的Pro版本, 可以为国家统计局办公室做一些有用的工作, 尽管与独立的图像相比, 功能性受到很大限制。Google Earth的优势包括就一些低级任务而言, 成本低, 易于使用。大多数图像的空间分辨率为15米, 可以让点查区划界者在一定程度上看清地貌, 但无法清点住宅单位。Google Earth的弱点包括达不到点查区划界所需的分辨率、难以将图像输入地理信息系统程序, 需要高速的因特网连接以下载图像, 还有元数据和真实性的问题。其他在线卫星数据来源, 例如通过环境系统研究所(ESRI)的免费ArcGIS Explorer应用软件获得的数据, 可以避开与Google Earth的一些问题, 允许将图像直接输入地理信息系统项目。不过可用图像的空间分辨率对一些人口普查应用来说仍然是不够的。

4. 63. 高分辨率卫星图像所能显示的地理细节与根据航摄照片创建的数字化正射地图相似。一个重要的复杂情况是比起时间安排灵活的低空飞行的飞机来, 卫星较难获得没有云层遮盖的图像。借助没有云层遮盖的高分辨率图像可以清点住宅单位、估计人口数以及划分点查区边界。航摄照片常常是特别拍摄的, 可能更适合详细的调查和制图项目。新的数字化航摄照片越来越受到人们的欢迎, 可能提供甚至比高分辨率卫星图像更高的图像质量(见下文4.73-4.89段)。

## 5. 将遥感数据用于人口分析

4. 64. 遥感技术有可能查明快速增长或变化的地区, 使得国家统计局办公室可以把资源集中在最需要的地方。利用遥感进行的人口分析还在起步阶段, 但正在快速取得成果。美国国家研究委员会的一份报告(NRC, 2007年)指出用于人口估计的替代方法对于人道主义反应来说仍不够可靠。高空间分辨率的传感器, 例如Ikonos和Quickbird, 没有Landsat所具有的档案数据的深度, 并且覆盖大片地区可能十分昂贵。不过, 正如所论述的, Landsat用于估计人口规模是有问题的, 并且其传感器方面的问题长期存在, 此外还有投资变异性。在不久

的将来，研究人员有可能可以利用其他高分辨率数据(例如SPOT)替换Landsat的应用软件。

4. 65. 可以利用卫星图像来估计的人口特征包括居住单位的计算、城市化土地面积(定居点范围)的衡量以及土地覆盖/土地使用的估计，这些可以替代居住范围和人口密度(Jensen和Cowen, 1999年)。在一些容易受灾地区，与卫星图像相比，航摄照片的好处是能够收集云层下方的景象。雷达在人口分析中还需发挥重要作用，尽管它有一个好处，就是能穿过云层。

4. 66. “变化情况探查”法可用来从空间的角度衡量人口的改变，尤其是查明快速增长的地区，即利用两张或两张以上同一个地方的图像，这些图像间隔的时间在五年或五年以上。为了对城市化范围进行量化，分析人员利用一种强行归类技术来给每张图像归类，这样每个像素就被认为是城市或非城市，这取决于光谱信号。随后通过覆盖图像和衡量居民区的增长，可计算土地覆盖层的变化。

4. 67. Antos援引的一个案例研究(Yankson, 2004年)使用了1985、1991和2002年Landsat TM的数据来计算这一期间阿克拉地区面积的年度增长率。Yankson发现1984年至1991年期间，阿克拉每年增长约10平方公里。1991年至2001年每年增长25平方公里。这一研究只衡量地区面积的增长。为了了解可能与发展有关的内部重大变化，必须采取一种软归类的方法，衡量作为一个持续变量的城市发展。分析人员不是把一个像素归为城市或非城市，而是可以利用其他的亚像素的数据来划分每个像素，用一个百分比来表示其城市化程度。这里赞成在利用地理信息系统来开展人口普查地图绘制活动时，使用一种持续的方法。

4. 68. 健康研究使用了遥感图像，以便重点关注在诸如疾病流行情况等现象中城市内部的差别。Castro(2004年)使用了航摄照片和社区多边形来查明有可能发生疟疾的地点。划定非正规定居点和社区的探查范围可以以定居点特殊信号的模式识别为基础，这些信号可能包括密集的屋顶、植物增长速度慢、普遍肮脏的大街小巷。一般说来，非正规社区一般有最少的纹理，就是说明亮度变化不大，密封的表面高度集中(Weeks, 2007年)。

4. 69. 最后，遥感数据的最先进的人口应用软件利用基于对象而不是像素的分析。Pellika(2006年)阐明了使用一种自动化的软件来把高分辨率的图像分割成大小、形状和色彩相似的区域，然后把每个区域标记为一个特殊的表面类型，例如“屋顶”。这样一种技术随后可以将一种特殊类型的屋顶全都归在一起，并标记为一个非正规定居点的一部分。Pellika将这一技术运用于几个案例研究，能够达到97%的准确率。基于对象的图像分析(OBIA)的缺点是虽然是自动化的，但仍要耗费大量时间；此外，它是针对地点的，依赖于拥有详细的高分辨率数据。目前，分析需要有昂贵的软件和专门技术，这些常常是大多数国家统计办公室工作人员所不具备的。

4. 70. 正如航摄照片一样，卫星图像的购买——尽管一般比航摄照片调查便宜——仍然可能十分昂贵。如有可能，应通过与其他机构的共担费用安排

来获取高分辨率的卫星数据，或者在地图覆盖面不足的地区有选择性地使用这些数据。

## 6. 卫星遥感数据的优缺点

### 4.71. 卫星遥感数据的优点是：

- 最新覆盖大片地区，较低空间分辨率的图像成本较低。
- 高空间分辨率图像能够覆盖一定的区域，其详细程度足以进行点查区划界，条件是划界地区的人口估计数存在。
- 根据图像可以对无法达到的地区绘制地图。
- 可以利用图像对实地核查工作进行独立检查。
- 图像有多种用途，一旦购买，可以用于其他地方。
- 在线的图像资源可以以很低的费用(或者免费)提供一定的功能性。
- 可以更新广大农村地区的地形图，例如，对原有地图中缺少的新居民点或村庄予以确认。

### 4.72. 缺点是：

- 许多系统，尤其是低成本系统的空间分辨率对于人口普查应用来说是不够的。
- 在光学传感器的情况下，云层和植被会限制图像的判读。
- 在发展中世界环境下地理特征的对比不够鲜明时，将很难区分，例如，很难区分农村地区的土路与传统的建筑材料。
- 图像处理需要大量专门知识，国家统计局办公室可能不具备这样的知识。在国家统计办公室面临的其他人力资源挑战中，这些机构可能决定利用其他机构的专门知识以执行本机构的绘图任务。利用这样一种机构合作的方式，可以分担费用，分享专门知识。

## D. 航空摄影

### 1. 航空摄影概述

4.73. 即使在高分辨率卫星图像出现的情况下，航摄照片对于需要高准确度和快速完成任务的地图绘制来说仍然有用。航摄照片类似于地图和卫星图像，因为拍摄到的是地表面各种特征的俯视图像。航摄照片与地图的不同之处在于，它们仅能表现地面上实际看到的特征，当然也没有人为的边界、专题信息以及注解。如果不作进一步处理，航摄照片达不到地图所具有的几何准确度。摄影机的角度和地形的变化使航摄照片的图像产生失真。因此，必须另外进行处理，才能制做出所谓的正射影像地图，这种地图兼有地形图的几何准确度和照片的高清晰度(见方框IV.2)。

4. 74. 摄影测量学是一门通过摄影图像获取测量结果的科学，用于制作和更新基本地形图，进行农业和土壤调查，并用于支助许多方面的城市和地区规划工作。人口普查项目还经常使用航空摄影调查，以便迅速为没有最新地图或难以采用传统实地方法进行调查的地区制作地图。在人口普查开始前不久进行的航空调查，将为在尽可能短的时间内为点查区的划界提供最完善的依据。

4. 75. 航空摄影在飞机发明后不久便一直用于制图。早期的航空摄影使用标准摄影机。但不久，定制的摄影机系统问世后将几何失真现象减到最少，并且安装在特别改造的飞机上，因此，摄影机系统能够通过飞机地板上的小孔垂直拍摄地面。判读航摄照片的设备以及将从此类照片抽取的信息转换成地图的设备很快发展得非常先进。例如，判读立体图像对是制作等高线地图的主要方法。

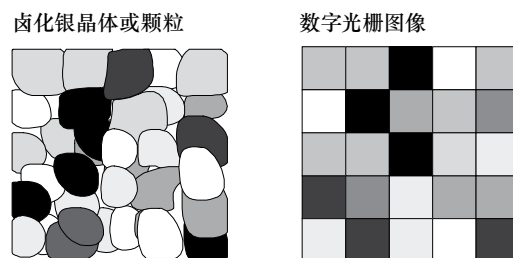
4. 76. 航空摄影是通过在低空飞行的飞机上安装的专用摄影机实现的。摄影机将图像拍摄在摄影胶片上或者数字化载体上。与数字传感器系统相比，传统上，胶片的分辨率要高得多(即区分细小物体的能力)，不过数字成像领域近来的发展改变了这种情况。根据航空勘查公司MJ Harden提供的文献，达到最新技术发展水平的传感器可以捕获12-bit的图像，地面分辨率小到每个图像像素1.5英寸，有4 096种灰度，相比之下，胶片有256种灰度。南非公司Rob Wooding和Associates将数字化航摄图像与1米的卫星图像进行对比，发现航摄图像不但费用低，而且更准确。这取决于所飞越的地区的位置，因此需要根据实际情况评估需求。

4. 77. 传统上，航空摄影项目的最终产品是地面某个区域的洗印照片。根据设计，航空摄影的照片，其重叠部分要在30%到60%之间。摄影绘图人员对这些照片进行泛化整理后，能够制做出覆盖整个地区的无缝镶嵌图。洗印出的航摄镶嵌图与地图的使用方法一样。这些图片可加注解，也可以供实地工作参考，此外能够通过地理特征的数字化处理创建或补充地理数据库。

4. 78. 数字化图像处理领域的最新发展改变了将航摄图像转变为有用的产品的模式。在模拟系统中，照片通常是一种中间产品。最普通的方法是将底片转换为照片，利用一种分辨率极高的扫描设备对照片进行扫描。结果就得到了可以在一台计算机上进行显示和进一步处理的数字化图像。照片与数字化图像之间的细微不同见图IV. 6。例如，黑白摄影胶片上涂有一层明胶，其中嵌有

图IV.6

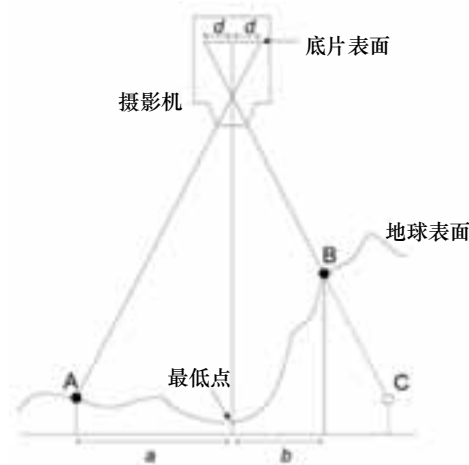
#### 摄影胶片与扫描图像的对比



方框IV.2  
数字化正射影像地图的开发

为了制作地图式的数字化正射影像，在图片上造成的失真需要予以消除，这些失真来自摄影机的角度与地形的变化。由于地形变动而造成的失真可参见图IV.7的说明(根据Jones, 1997年)。摄影照片从本质上来说，实际上是地球表面的一个透视投影的结果。点B与点A相比处于比较高的位置上。在实际情况中，B点的位置从水平方向的最低点算起，距离为b，这个最低点处于摄影机镜头透视中心垂直以下的位置上。可是，反映到摄影机底片上的透视投影结果却给人们一种视觉上的误导：B好像是位于点C，因此，投影到底片中心的距离为d，与点A的投影是相同的。

图IV.7  
地形所造成的失真



为了纠正航摄照片上的失真，必须知道地面上每一个点的标高。标高可以通过一对对的立体航摄照片确定。这些航摄照片覆盖了地面上大致相同的区域，但两幅照片在拍摄时稍微错开了一小段的距离。立体测绘分析仪提供了相应的功能，操作人员可以借助它把立体对的图像正确标注，并将其中的三维空间特征位置抽取出来。目前一流水平的软拷贝地图绘制系统能够具有非常高的自动化功能，可对图像进行对准并消除失真。在校正过程中可考虑所有相关参数，例如，空中摄影机的倾斜、镜头的失真等。由此，操作人员可以从航摄照片中正确地抽取地理基准的数字化数据。输出的成品包括：从航摄照片直接生成地理信息系统向量数据，生成可以显示地形变化的网线框架式地图，或者生成数字化标高模型，后者是一种与航摄照片相对应的光栅图像，其中的每一个像素值代表地面上相应点的高度。数字化标高模型虽然在人口普查制图应用方面用处不大，但这类数据集在环境和自然资源应用、特别是水文方面用处很大。

在适当的地理基准系统中完成了上述对准过程并消除失真后，原始的航摄照片已经被转换成数字化的正射影像地图。它们通常按1:2 000至1 20 000的地图比例制成，这取决于拍摄高度和处理工作。一幅幅相互连接的正射影像可以通过数字化方式组合到一起，创建出一套无缝整体图像数据库，用于整个城市、地区或整个国家。制图人员通过屏幕数字化处理从正射影像中抽取或勾画出地理特征。或者，直接将其作为背景图，供现有的地理信息系统数据层使用。

细小的光敏卤化银晶体。这些晶体或颗粒呈不规则状，且大小不等。相反，扫描图像是规则的像素排列(图像的组成部分)。

## 2. 航摄照片在人口普查制图方面的应用

4. 79. 正射影像地图非常适合用于居住单位计数和人口估算。借助于航摄照片进行的房屋和人口统计有时被称为屋顶测算。在农村地区，居民点在航摄照片上可以很清楚地得以区分，房屋基本上处于分散状态，居住单位的数目可以相当容易地加以确定。如果有了每个家庭人口平均数的可靠估算值，即可得出对于人口普查目的来说足够准确的人口估算值。在市区，房屋可能非常紧密地靠在一起，住在多层楼房里的家庭数目也很难确定。即便如此，如果对该区域有所了解或调查员受过一定训练，仍然有可能得到足够准确的人口数量估算值。然后，人口普查员可以划定点查区的边界，使每个点查区内包含一定数量的住宅单元。由于对正射图实施了正确的地理基准工作，因此点查区的划分也可以借助已知参数按照合适的地图投影方法进行标注。这意味着，为了使数字化边界与其他地理信息系统数据保持兼容而不得不做的可能繁琐的地理基准工作就没有必要进行了。

4. 80. 航摄照片的解释通常基于人的视觉解释。因此，人口普查制图员没有必要接受高级的图像处理技术方面的专门培训。点查区的边界可基于航摄照片进行划分，为点查人员提供地理基准的附加地理特征也可从航摄照片中抽取。这种特征也可在计算机上通过交互方式予以界定，只要用鼠标或其他类似定位装置即可完成。还有一种方法是，人口普查员可以把图片打印出来，在透明的(醋酸纤维或聚酯纤维)塑料胶片上对特征进行追踪。然后也可以扫描这种图片，进行向量化处理。这种处理方法需要一个额外步骤和更多的材料，但经过这种处理后，通常能提高成品输出的准确度(另见有关数字化和扫描问题的部分)。

4. 81. 必须强调的是，不管是基于卫星遥感还是航摄照片的点查区边界文件如果没有适当的参照数据——例如陆标和街道名称——都是没有意义的，因为点查员可能无法确定自己的位置。

## 3. 航空摄影的实施与机构问题

4. 82. 建立数字化的正射影像需要在摄影测绘方面具有相当的专业技能，而一般人口普查机构不具备这些条件。因此，人口普查机构需要与其他国家单位签订合作协议，这些部门大多是地图测绘单位或空军的勘测单位。此外，还有一种办法，就是把工作包给一家商业航测绘图公司。现在，已有几家国际性的地图绘制公司，它们可以提供飞机、摄影机和数据处理设备。

4. 83. 然而，此类服务的价格不菲。但好在航摄照片可用于多种用途，例如服务设施的规划、城市地图更新和地籍项目等。如果把费用分摊给有意加入的各政府部门，或者，甚至吸收一些私营部门，就可以大大降低人口普查机构所要承担的费用。由于资源所限，有时不可能制作覆盖全国的航摄照片，但可为特定地区制作。例如中国香港统计局就曾利用航空摄影的手段估算在船上生活的居民人数(见Netherlands Interdisciplinary Demographic Institute, 1996年)。

这个实例说明，可以利用这项技术计算那些难以直接进行点查的人口。另外类似的例子还有游牧人口或难民群、飞速发展的市区，或者有些季节难以到达的地区。

4.84. 如上所述，开发正射地图需要相当的技术力量和专业设备。相比之下，使用正射影像地图并不要求接受很多严格培训，尽管应该与整体的人口普查计划兼容。例如说，一个城市的数据库可能只是一个便携式存储媒介如DVD上几张可以拼接的图像，通过一个标准的地理信息系统或桌面地图绘制软件可以将其无缝连接在一起并显示到屏幕上。数字化正射影像地图可以按标准图形格式(如TIFF标记图像文件格式)存储。因此，用户并不需要配备专用的图像处理软件。事实上，任何一种图形软件包都可以用于从这些图像中抽取所需特征，不过这样操作时，地理基准信息会有所损失。所丢失的信息通常包括数字图像的尺寸与真实世界坐标信息，它们一般存储在一个很小的首标文件内。借助于这些信息，大多数桌面制图软件包都可以将这些图像标注到任何地理信息系统数据集中，只要这些数据集存储在相同的地理基准系统内。

4.85. 正射影像地图还可作为背景，在上面显示全球定位系统所收集的点的位置，或其他数字化特征，如健康设施和运输网络等。过去，除了制作点查区地图之外，还可将数字化正射图打印出来，发给点查员，帮助确定指定给他们的边界。不过，今天，遥感图像更有可能成为一个地理信息系统项目的组成部分而不是作为一个单独的地图包括在内。

4.86. 有一个问题妨碍了这种技术在人口普查办公室内的应用，即，对于大面积、高分辨率的数字化正射影像地图而言，工作中所涉及的数据量过大。对一个人口普查办公室来说，为了便于工作，最好能获得分辨率相对较低的数字化航摄照片，只要其显示的内容能够满足人口普查工作的需要即可，以方便处理和存储。数字化正射图常常具有非常高的分辨率，每一像素的地面尺寸可以达到厘米一级(通常为5-30厘米)。通过对数字化正射图重新取样，使每个像素的尺寸限制在0.5-2米范围内，这样的分辨率已足以满足在市区范围对点查区进行划界了。

4.87. 未来的航空摄影工作将是全数字化的过程，因此无需制作中间环节的印刷图片。使用飞行全球定位系统控制和数字分幅摄影机的系统已经投入使用。数字分幅摄影机采用电荷耦合器件(CCD)，可以生成 $9\ 216 \times 9\ 216$ 像素的图像，其实物定位准确度可达1至4厘米。由于取消了一些中间步骤，如照片印制和随后的扫描工作，使用这项技术比使用传统的地图绘制技术速度快，价格便宜。正如计算机的处理速度一样，数字相机的分辨率肯定也会不断提高。在不久的将来，准确、实时和全数字化的航摄地图制作技术一定会取代传统的航摄技术。

#### 4. 航摄照片的优缺点

4.88. 航摄照片的优点如下：



- 航摄照片可以提供大量细节，并可进行视觉解释。可同时显示许多地理特征信息，如道路、河流以及建筑物等。
- 数据收集迅速，地图数据的加工制作快于地面测绘。与很少更新的地图相比，最新的航摄照片作为人口普查地图绘制工作的基础更为可靠。
- 对于交通不便或野外操作十分困难或十分危险的地区，可利用航摄照片制作这些地区的地图。
- 利用航摄照片制作地形图比利用传统的测绘技术的价格低。不过，由于人口普查地图的准确度要求比地形图所要求的准确度低，如果仅为人口普查目的而投入大量投资则没有必要。
- 印制出来的航摄照片对野外作业十分有用，它是一幅“更大的画面”。野外作业人员可以从更广阔的视野看到他们四周的地形。在显示地理信息系统数据集时，数字化航摄照片可以作为背景使用。

#### 4.89. 缺点如下：

- 航摄照片的处理需要昂贵的设备和专门技能，因此，人口普查办公室需要与其他机构合作，以获取正射照片或者以其他方式依靠外部的支持。
- 航摄照片也许拥有版权，发布权利有限。
- 航摄照片还需要地理特征名称信息，这些信息只能从过去的地图中抽取；航摄照片无法完全摆脱对野外作业的依赖。航摄照片对于偏远的地区也许是不够的，尽管它可以为难以点查的地区提供背景。
- 当地理特征隐藏在茂密的植被或云层之下时，或当对比度不够，难以区分紧靠在一起的地理特征时(例如难以区分天然材料建造的住宅与周围的地面环境)，航摄照片的判读将发生困难。
- 数字化航摄照片包括大量的数字化数据，必须使用高性能计算机显示和进一步处理。

## E. 概要和结论

4.90. 第四章审查了在人口普查实地工作中如何利用工具，例如全球定位系统和遥感技术，以及将这些工具与地面工作相结合。利用这些地理空间工具可以让国家统计局办公室将工作的重点放在国内快速变化的地区上。到了人口普查过程的这一点上，国家统计局办公室已经完成了点查区的地理数据库。接下来的步骤是设计并从地理数据库中打印和分发地图，供实际点查使用。

## 第五章

### 在人口普查期间使用 地理数据库(地图)

#### A. 引言：人口普查的点查期间 使用地理空间工具

5.1. 在人口普查之前和之中，地图被用于各种规划目的。使用地图的活动包括将点查员分配到各个地方；查明道路崎岖的或无法达到的地区；为实地工作人员的交通和用品管理后勤工作；确定难以点查的人口和集体生活区的位置；在多个层面上划定行政区边界；监督人口普查进度；以及绘制定位地图。

5.2. 本章将讨论绘制这样的人口普查地图的过程，重点是如何利用地理数据库的多功能性来及时和有序地将正确的信息送到点查员手中。更广泛地说，本章阐明了地理空间技术如何能在点查阶段支持人口普查活动。本章涉及的主题包括地图编制、为点查员和监督员确定相关的层以及打印和分发的基本要素。

5.3. 总的说来，本章将对点查采取一种项目管理方式，即以计划为中心。如果对后勤工作做了详细规划，错误就会降至最低，窝工现象也可以避免。人口普查是一种地域性的活动，在人口普查中，国家被划分为各个操作单位，在这些操作单位中可以进行详细调查。数字化点查员地图的优点是可以针对国内进行人口普查的特殊背景进行修改。建立数字化的点查区使得国家统计局办公室可以拥有一份活的文件，该文件既利用了以前的人口普查工作，又增加了遥感技术和全球定位系统的价值。

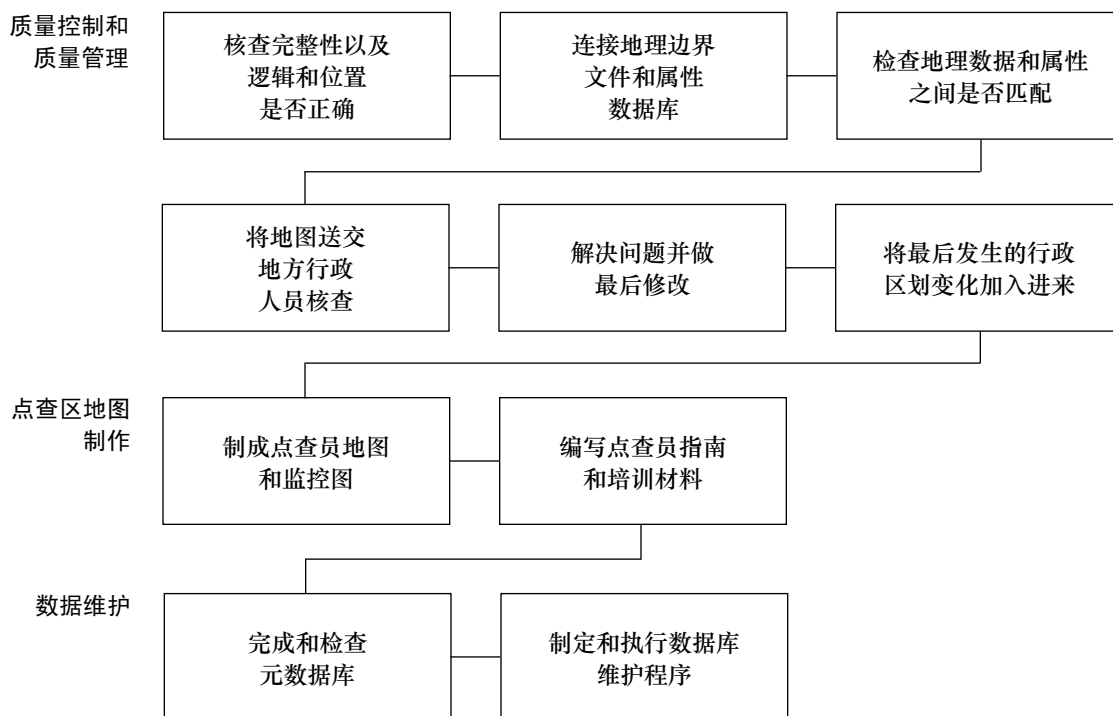
5.4. 到了人口普查进程的这一阶段，对以前的人口普查和所有实地工作进行地图编制已经以地理基准化格式告一段落。按照第四章中介绍的变化情况探查方式，人口普查规划员已经查明需要最多关注的地区。如果以正确的方式开展了工作，国家统计局办公室在这个时候已经是万事俱备，只欠计数。这里的观点是经过全面编辑和更新的数据库将成为点查员带往实地的地图的基础。随后，为了进行点查，任务是从点查区数据库中输出地图，交到点查员手中。

#### B. 质量保证、点查区地图制作和数据库维护

##### 1. 概述

5.5. 人口普查数据的准确度和完整性基本上取决于点查员使用的地理底图的质量。除了持续的质量控制程序和数据转换期间的质量改进，在将点查区

图V.1  
质量保证、输出成品和数据库维护阶段



地图分发给点查员之前的最后一个步骤是全面检查所有地图产品。这也包括由地方管理人员核实行政边界是否正确。在制成最终产品之前，任何剩余的问题和不一致的地方都必须解决。

5.6. 在这个时候，人口普查的实地办公室应该已经建立。人口普查组织结构的集中程度将对实地分发的点查区地图和小组负责人地图产生直接影响。在实地办公室，可以展示一种根据点查区地图编制的比例更小的地图，以显示进度。实地办公室的其他规划工作包括估计工作量和旅行费用；安排分发和接收材料；确定有麻烦的地方；以及安排办公室人员去实地视察。

5.7. 从概念上说，制成点查区地图是一件简单的事，只要数字化地理数据库的质量令人满意。这个步骤主要对后勤工作构成挑战，因为必须分发数千张地图，以及地图阅读说明和其他指示。

## 2. 草图制作和质量保证程序

5.8. 以下各节讨论为点查工作制作人口普查地图的全过程。步骤包括地理数据库的集成、地图特征的编制、地方当局的核查以及地图的打印和分发。

## 3. 边界线和属性文件的匹配以及总图的打印

5.9. 在最终的地图设计和打印准备工作中，如果边界数据集和地理属性文件尚未集成到一个相容的数据库中，则需要对其进行匹配。这一过程也包括

对边界数据和地理属性数据匹配是否正确的检查。如果两者均正确,针对地理属性文件里的每一个记录,就应该至少拥有一个地图特征(一个点、一条线或一个多边形)。如果情况并非如此,则意味着地理数据库有错误,即遗漏了一个点查区,或地理属性表含有重复记录或错误记录。假如一个属性记录有两个或更多的多边形,质量保证人员就必须查明有无违反操作规范的行为。

5.10. 那些起草点查区地图的人应该铭记,点查区地图必须完全覆盖国家领土。将点查区边界看作是围绕点查员任务区的“篱笆”有助于确保全面覆盖。当点查区边界使用街道和公路的中心线时,错误可以降至最低,因为这样一来点查员可以在街道或公路的一边详细调查居住单位,而把另一边留给相邻点查区的点查员。在第三章和第四章中可以找到关于点查区划界的其他规则。

5.11. 地理数据和属性信息正确匹配以后,需要给地图添加标记,选择的地图符号必须能识别底图上的地理特征(另见第六章有关通过地图传播信息的内容)。标记工作可以利用地理信息系统软件包或更专业的制图设计软件,以交互手工、半自动或自动的方式进行。在一个非常大的人口普查地图绘制项目里,为特征添加标记是耗时极多和烦闷无比的工作。尤其是当点查区地图的设计相当复杂时(例如,综合多张数字化地图层以制成每一张点查区地图),按人工工时和计算机资源计算,正确的标记布置所需的资源可能非常大。

5.12. 大多数地理信息系统和台式绘图系统可提供自动标记布置功能。用户只需指定用于标记的地理信息系统数据库属性表里的属性域即可,例如,指定街名或建筑物的标识符。系统会利用一些简单的规则在每一个特征上或其附近安放标记。标记的大小通常由用户决定,而且,如果特征过于靠近,也可决定是否应将标记叠放在一起。然而,除了最简单的情况外,仍然需要手工进行一些标记的修正工作。

5.13. 为了完成庞大的点查区制图项目,人口普查办公室可能应该考虑购买一个名称安放专用软件包,如Maplex。这些程序能够进行较复杂的计算,确保遵守最重要的标记安放规则,规则包括:

- 标记之间没有重叠或重叠量最小。
- 特征和标记之间没有重叠或重叠量最小。
- 标记能明确说明特征(不模棱两可)。
- 总体图像效果美观,例如,字体和字号都很合适。

5.14. 软件包为安放标记提供了一些试探性规则作为基础,用户可根据专门用途对其进行修改。用户可以省去个别注释数据层里的,为特定的地理信息系统数据层设计的标记,而将其覆盖在地理特征数据层上。

#### 4. 质量保证

5.15. 虽然在计算机屏幕上可以对统一性进行交互式检查,最终的质量保证工作最好还是通过打印出来的硬拷贝地图来进行。因此,应该制作大版面

的地图，包含最终点查区地图上所有应该出现的信息。这些地图是为了最终质量保证和检验而制作的，应该由行政单位组织实施。如果打印出来的地图和最终点查区地图的比例相同，每个区就需要若干张这样的地图。

5.16. 质量保证系指将数字化地图数据库产品发布给人口普查工作之前的最后一次检查。质量保证和质量控制类似，在第三章中已经阐述质量控制。有些检查将针对所有产品进行，而更复杂和费时的检查将采用适当的验收抽样方法针对产品的某一子集进行。

5.17. 数据转换过程的质量控制集中在边界和坐标的拓扑以及位置是否正确上。务必保证已经数字化处理和分别存储的边界无缝衔接。例如，地区图若存储在不同的数字化地图文件里，相邻地区之间的边界必须完全相同。质量保证强调的是，最终的地图产品要适合于点查工作。因此要在几个方面对数据库完整性进行检验，这些问题将在下文论述。质量保证不是微不足道的琐事，要耗费大量的时间和资源，人口普查办公室应相应安排好日程和预算，不过如果做得出色，最终会使人口普查更为准确。

5.18. 人口普查制图人员进行的检验，应涉及以下验收标准检查：

- **易读性。**地图上的所有注释必须清晰易读。有时，一张地图上画了过多的特征，导致街名和其他文本信息难以辨认。有些非关键文本标记可以省略，以提高地图的清晰程度。同样，必须清楚表明每个文本标记指代的特征。有时，可能需要通过箭头指明对应关系。
- **数据层。**将数据层画到地图上时，先后顺序关系重大，因为，位于上面的图层可能会盖住较低的地理数据图层上的重要特征。
- **地图比例。**例如，一张很大的、但包含一个相对较小的拥挤区域的点查区地图可能需要一张插页或者一张局部图，以保证所有细节都能识别。
- **来源和版权信息。**为了制成点查区地图，需要建立数字化数据库，如果为此使用了具有专利的数据源，应该将其一一列出，包括制成地图的日期，这样点查员可以确定是否是最新日期的地图，是否需要进一步的实地修正。

## 5. 地方当局的核查和最终的行政单位检查

5.19. 打印出来的点查区地图应呈送地方当局审查，这是至关重要的统一性检查。需要让人口普查机关内外的地方行政官员确认，所有较大城镇的居民点和分区是否已经纳入到地理数据库中。让地方当局参与这个过程的好处在于，可以由熟悉当地情况的人将地图复审一遍。在使用几种语言或方言的国家里，名称和拼写的习惯可能不同。地图经过地方官员审核，就可以减少由当地雇用的点查员判读地图时出现的差错。

5.20. 确认点查区地图上的行政区划边界也是审查过程的一部分。这些边界常常改变，经常增加新的州、省和区，无论是否开展人口普查。有时，这可能给国家统计办公室带来麻烦，这些机构需要为这些行政区进行摘要统计。

管理行政边界的程序在第三章中做了论述。在理想的情况下，在进行人口普查之前数月应通过政府命令冻结行政边界。这就为人口普查期间提供了一个稳定的参考框架。这一阶段所处的边界结构是将其制成人口普查表格的结构。临近点查时处理这一问题的选择包括：

- (a) 一个办法是，在人口普查前持续追踪行政区边界划分的变化。变化一旦发生，立即将其反映到数字化地图数据库中。这样做的结果是，在进行点查时，边界正好就是当前的实际边界。然而，持续不断地监视变化并修改边界数据库需要占用额外的资源。国家统计局办公室应考虑使用空间数据库引擎，记录各行政边界确立的日期；
- (b) 在有些国家，边界的改变都会事先公布。这样，人口普查制图机构就可将这些地区的人口普查制图工作安排在较后的时间进行；
- (c) 最后一个办法是，让人口普查制图单位规定一个冻结日期，过一段时间以后再修改全部边界，很可能要到人口普查结束之后才能进行。如果修改后的行政区划边界穿越了现有的点查区或其他边界，这些单位的住户调查表则必须按正确的单位重新分配。这意味着在点查后又增加了一个步骤，人口普查结果的发布也因此被推迟。

## 6. 点查区地图的制作(包括地图打印)

5.21. 对所有底图和点查区划界完成了审查和质量保证工作后，人口普查制图人员就要打印最终的监管图和点查区地图。这一工作可以在国家统计局办公室的主要地点或者区域的实地办公室进行，这取决于国家统计局办公室的组织结构。监管图将显示几个点查区，并以较小的(即更泛化的)制图比例打印。为具体点查区规划平面布置的方法与人口普查制图法的剪切步骤类似。点查区地图必须简单明了，因为使用它们的是看图经验有限的点查员。另一方面，它们又必须含有足够的信息，使操作者容易定位。点查区地图应具备以下信息：

- (a) 整个点查区域，边界线要明确；
- (b) 要显示出邻接区(即周边区)的一些部分，以便定位；
- (c) 任何便于在点查区范围内定位的人口普查制图数据库中的地理和文本信息，利用关于符号化的标准的地图协议(例如，使用虚线来表示小路，用蓝色表示水，等等)。点查区地图的特征应包括：
  - (i) 街道和公路；
  - (ii) 建筑物；
  - (iii) 地标；
  - (iv) 水文特征；
  - (v) 其他显著特征或相关特征，可能包括地形、水体等；

- (vi) 统一的地图图例或地图关键字，包括行政区和点查区的准确名称和编码、指北箭头、比例尺图解以及说明用于地理特征的符号的图例。

5.22. 图V.2显示了一张假设的城市点查区图的各组成部分。所有的特征都存储在空间参照系相同的各自的地图层里，或者形成图形模板。主要的组成部分为街道网络、建筑物和点查区边界层。此外，注释、符号、标记和建筑物的数量均存储在各自的数据层中，需要时也可动态添加。最后一个组成部分是模板，由图标边线(一个方框表示地图的外延)和所有点查区共同使用的图例组成。图V.3显示的是一张完整的点查区图，所有的组成部分都叠加在一张图上。比起这张样图来，点查区图包含的信息可以多也可以少，视人口普查活动的范围大小和点查区的复杂程度而定。

5.23. 在没有对即将进行开展的人口普查进行全面数字化的许多国家，点查区的地图设计可能比图V.3简单。例如，不使用向量格式的完全集成式数字底图，而把地形图的光栅图像作为点查区边界线的底图使用。在有些情况下，地图的特征可能更为泛化，例如，就用中心线表示街道，用多边形表示整个城市街区而不是单幢房屋。

5.24. 必须对打印的点查区地图的格式和颜色做出决定。(打印机选择标准，见第六章)倘若激光打印机能提供高分辨率，点查区图应尽可能在A3纸(420x297mm, 两张A4纸大小)或11x14英寸的纸张上打印。与较大版面的打印机或绘图机相比，标准尺寸的打印机的好处是成本较低而输出速度较高。如果按每页计算，包括墨水或墨粉的费用，激光打印机比喷墨打印机便宜很多。由于需要制作成千上万张点查区图，这些都是应该考虑的重要因素。问题会发生在那些点查区非常大的地区，这些点查区包含了一些既小又拥挤的小区。对于这些地区，必须打印版面较大的地图，或者在地图设计里加进插页，进一步显示点查区的稠密部分。

5.25. 为了在开始点查前弄清楚纸样地图库，国家统计局应考虑给地图打上条形码。条形码是一种容易实施的技术，利用一个激光阅读器和简单的数据库软件就行了。条形码不包含说明性数据，只是计算机用来查找相关记录的随机的参考号码。条形码阅读器感觉反射光的明暗，将其转换为一种电子信号(黑色高，白色低)。许多条形码阅读器如今带一个USB界面，解码器或者内置在装置中，或者是单独的。国家统计局可以利用预先打印的条形码标签或者使用为每一幅制成的地图建立一个条形码的软件(更多信息见[www.barcodehq.com/primer](http://www.barcodehq.com/primer))。

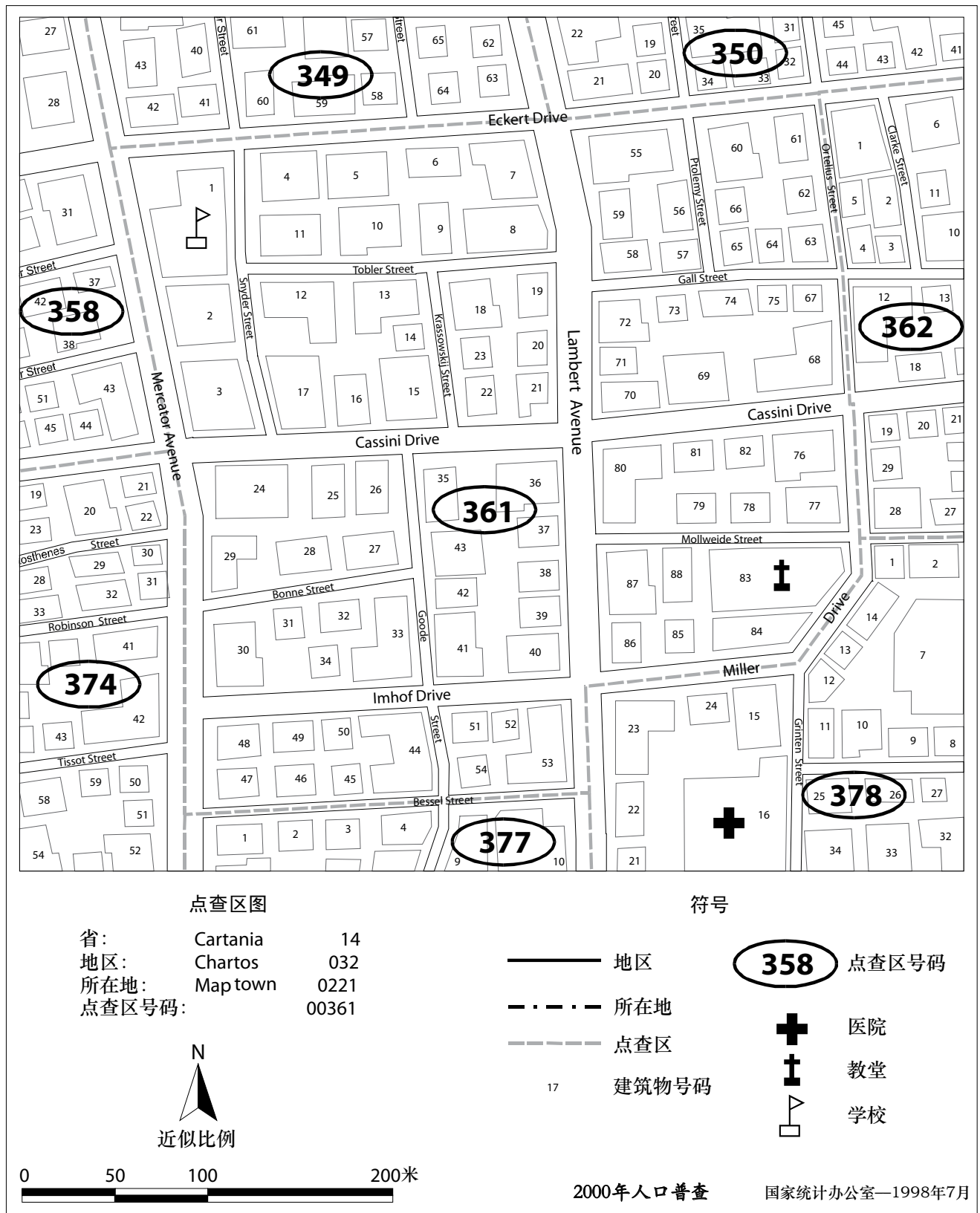
5.26. 一幅设计良好的点查区图，采用黑白图即可满足工作需要。虽然彩色打印机的价格下降了，但与黑白打印机相比打印速度太慢，效率方面有局限，而且消耗材料相当昂贵。良好的黑白地图可以照像复制而不致丢失信息，这样，地方上的工作人员在需要添加点查区图复制件时，可以自行完成。不过，如果资源条件允许，使用彩色可以使地图版面更加醒目。例如，在地图上的点查区边界处就可以利用明亮的颜色线条予以标出。同样的效果也可以使用

图V.2  
数字化点查区地图的组成例子





图V.3  
城市点查区地图的例子



明亮颜色的荧光笔在黑白地图上沿点查区边界标出而达到(关于颜色使用专题的更多讨论见附件五)。

5.27. 如果可能的话, 国家统计局办公室的官员应该考虑对容易读的地图进行可用性测试。随后可以在用户的实际体验而不是一个室内设计小组自己的判断基础上做出设计方面的决定, 例如地图大小、色彩和比例以及地图要素。

5.28. 每一幅点查区图要准备多个副本, 还要为中央人口普查测绘办公室提供保存用的副本。每一幅点查区图都要提供给当地人口普查机关、监督员和点查员, 总计约需4至5份副本。是以数字化的方式还是硬拷贝的方式发布地图的决定取决于人口普查活动的集中程度如何。如果地图处理工作集中在一个或少数几个人口普查机构内, 最好采取分发数字化地图文件的方法, 而不必发放硬拷贝地图。传送这种文件给地方人口普查办公室可以利用CD-ROM或DVD或在一个专门的ftp网站上通过互联网发送。地方办公室不需要直接操作地图处理软件, 只要把地图导出到一个常用格式的文件上即可(例如PDF格式), 或将其作为图形文件嵌入到一般的字处理格式的文件之中。这样的文件可在任何一种通用的计算机系统上打印出来。这种办法可使地方人口普查办公室按其所需制作点查区图的多个副本, 遇到丢失硬拷贝地图等情况, 可立即得到所需的文件。在各种选择中, 人口普查官员应选择最可靠和最具成本效益的方法。

5.29. 如果人口普查地理数据库本身一致, 结构完好, 点查区图的打印应能很快完成, 通常不需外包这项工作。打印点查区图并不要求使用高端地理信息系统软件包, 应用相对较低廉的桌面地图处理软件即可。运用软件中内置的宏语言, 其中有些过程可以自动进行。例如, 一个点查区列表可以配上点查区边界坐标(所谓的地图周边范围), 并标出坐标单位。软件将按指令检查该表, 把包括数据层的内容放入到预先准备的、标有图例和其他周边信息的模板中, 按要求打印出规定数量的副本。

## C. 在人口普查的点查期间使用 地理空间基础结构

5.30. 地理数据库对一次成功的人口普查所做贡献主要发生在实际点查工作开始之前和结束之后。不过, 在人口普查点查期间, 地理空间基础结构也可发挥作用, 支持后勤保障计划和监视普查进程。同时, 点查过程给人口普查办公室提供了一个对数字化人口普查数据库再一次进行质量控制的机会。以下对这两方面加以讨论。

### 1. 运用数字化地图作人口普查的后勤保障

5.31. 在人口普查过程中, 地图的使用出于多种目的。其中, 地理空间技术在规划点查的起步工作和后勤保障方面也能发挥积极作用。为操作区分配、确定实地办公室的位置, 以及安排实地工作人员和点查员的行程等都是地理数据库可以发挥作用的方面。如果将数字化地图用于这些目的, 而且国家统计局办公室选择不创建一个数字化的地理数据库, 人口普查制图单位则能利用一

个粗分辨率的地理空间数据库，其中有居民点、道路、河流和行政区划的小比例地图。这些地图的比例在1:50 000至1:500 000之间；1:1 000 000可能太宽泛，用处不大。在大多数情况下可从现有来源得到这些数据集，包括国家制图机构或勘查部门。国家统计局办公室可进一步开发自己的数据集或者利用地理信息系统软件附带的“现成的”数据集。

5.32. 很多地理信息系统软件包具有网络分析的功能，借助这项功能，规划人员可以确定沿某一道路网行进时的距离和费用。道路数据的质量因位置的不同而有很大差别，在小比例(即高度泛化的)数据的情况下应该小心。在城市化地区，规划到点查地点的旅行不是什么大问题，可是在农村地区，过长的距离和自然环境将造成旅行困难，增大实地活动的开支。在确定实地办公室的位置时，这也是一个要考虑的因素，实地办公室分管若干监督工作或小组负责人的所在区。选择实地办公室应使旅行时间最短，以便地区人口普查管理员发挥监督作用。通过地理信息系统软件程序能够统一观察各个地区，借以确定和显示各种可能的地区性分派任务的方案。

5.33. 将地理数据库用于后勤保障不如将其用于点查区划界那么重要。很多这类任务借助普通出版的地图同样能够很好地完成，尽管这么做降低了地理数据库的可用性。为此目的使用地理数据库可对旅行时间和距离估计得更准确，并且，人口普查人员可以迅速绘制出能够显示普查规划过程各个方面的地图。再者，为全国开发一套小型的粗分辨率的地理数据库，对于制作详细的地理基准普查数据库来说是一次很好的演练。

## 2. 监督人口普查工作的运作过程

5.34. 在人口普查期间和点查后立即进行的活动期间，总部工作人员将监督点查和数据汇编过程。通常，地区人口普查办公室对点查活动完成信息和最初结果进行编制。总部将收集这些信息，并评估进展顺利的地区和可能遇到问题的地区。

5.35. 有些国家实行一种所谓快速计算策略，即，将人口总数快速编制，然后同事先估计的数字加以比较。如果某些地区所报数字异常高或异常低，则需立即予以关注。传统上，这些评估是以表格形式编制的。如存在一个详细的数字化人口普查地图数据库，这些信息则可按地理区划予以显示。通过这种方法能够很容易查出有问题的地区。

5.36. 实际上，任何适宜的摘要统计都可以通过一个标准的关系型数据库系统进行编制。例如，一个用于说明报告区点查工作是否已经完成或每个地区有百分之多少的点查区已经完成工作的指标。人口普查人员可以定期地把各条信息连接到地理数据库，准备进行地图输出，供人口普查工作所有监督员进行评价。

5.37. 快速质量控制过程的关键在于快速的信息流动，从监督员到地区办公室，再从地区办公室到达总部。最快的信息交换办法是利用因特网。如果地方和地区的监督员有条件访问因特网，信息可通过网络，由一个受密码保护的数据库界面发送。

### 3. 点查员在人口普查期间利用地图的准则

5. 38. 针对点查员的培训方案应该包括基本的地图阅读和导航技术。一旦到了现场，点查员应在地面和地图上确定其点查区的位置。其次，点查员应确保使用手中的地图，尽管自己所处的位置似乎一目了然。确定地图的方向意味着把地图展开，这样地图的所有部分都能看到。点查员应该在一条主要街道的十字路口开始定向，在地图和地面找到地方，确定特征——房屋、礼拜堂、铁路、河流——并且对着地图上的符号检查这些特征。

5. 39. 点查员，在小组负责人的帮助下，应该围绕点查区确定旅行路线，尽量不要走回头路。可以把一个点查区划分为小区、街区或部分。一天的详细调查开始时，点查员应该在起始点做一个标记“X”，当天结束时，在终止点标上日期。点查员应该遵循“靠右”的规则，从起始点开始顺时针行进，这样点查的房屋就在点查员的右边，并且确定地图的方向，使其指向旅行的方向。在每一个拐角处，应重新确定地图的方向，使地图上的特征与地面上的特征相符。点查员应在地图上确定住宅单位，以便必要时进行回访。地面障碍，例如地形、水和土路，应该记在地图上，这样以后的点查员可以利用该信息。

### 4. 在点查期间更新和修正点查区地图

5. 40. 在准备点查区地图期间，即使执行一整套质量控制程序，许多地图很可能还会包含错误，有的还是严重错误。例如，在最初的实地工作期间，有些建筑和街道可能被忽略或者没有在地图上正确标注。由于为人口普查构建地理基础的实地工作需要点查前几个月甚至几年进行，新的建筑物和基础设施无法标记到后来点查员使用的地图中。

5. 41. 除了训练数据收集和基本的地图阅读技能之外，人口普查办公室还应告知点查员把点查期间发现的任何错误或疏漏标记到点查区地图上。在普查结束后，普查制图人员应该收集点查区地图，通过将修订纳入人口普查地理主数据库的既定程序，跟踪任何提议的修订(关于该主题的进一步讨论，见第六章)。这可能需要在数字化人口普查数据库中做相应的修改，或者可能需要另外进行实地检查或获取卫星图像或航摄照片数据来核实更改之处。这可能是一个非常复杂的过程，但应确保国家统计局拥有关于点查区的最新信息，这将减少未来人口普查或调查之前的绘图活动的工作量。不过，不应推迟发布人口普查结果。

## D. 概要和结论

5. 42. 第五章讨论的是创建点查所需的地图的过程，操作重点是讨论地理数据库时所遗漏的。包括的部分有点查员地图的处理流程、编制、为点查员和监督人确定相关的层、打印和分发基础、点查的项目管理方式以及出现拖延时的应急计划。

## 方框V.1

## 印度人口普查实地地图的制作

1. 在印度进行人口普查是一项巨大的工程，因为要清点和收集国内超过10亿人的信息。在确保全面覆盖国家地理区域，既不遗漏也不重复，制图起着关键的作用。
2. 几乎提前三年就开始了地图的筹备工作，地图上要显示各个邦、区、分区、村庄和城市的行政边界。从省政府和市政当局那里收集了指出管辖区变动的通告。对上一次人口普查时制作的地图进行了修订，指出了报告的改动之处，并且在用于人口普查前经过了指定当局的证实。管辖区方面发生的任何进一步改动都被收录了进来，直至为人口普查之目的冻结行政边界。人口普查组织有一个边界文件数据库，用于利用地理信息系统技术进行的更新活动。
3. 在人口点查前六至八个月进行的编制房屋清单的时候，划出了由列入清单的房屋组成的点查区，以便分配给大约两百万名点查员。由于没有可利用的数字格式的显示建筑物、房屋等详细布局的地理数据库，而从以前的人口普查获得的显示建筑物、房屋、公路和大的地标的布局的不用仪器而随手画出的地图可被用于农村地区。在城市地区，可以从地方当局得到的最新地图被用来划分由列入清单的房屋组成的点查区。根据在编制房屋清单的活动期间收集的信息，划出了新的点查区，覆盖地区的平均人口为500人至750人，供点查时使用。不用仪器而随手画出的点查区，即“概念地图”被点查员在人口点查期间使用，这些地图同样与其他机构共享，以便进行实地调查。
4. 尽管一直到市镇和乡村一级都有显示行政边界的数字化地图，但是对于市镇和乡村还没有显示建筑物、房屋、小巷、僻巷、道路网和大的地标的布局的详细的数字化地图。目前，已经开始着手建立大城市的详细的数字化地理数据库，该数据库将利用卫星图像。国内最重要的政府制图机构——印度测绘局——为此将在卫星图像的基础上，以数字化格式提供详细的行政区地图。提议进行特殊的实地调查，以收集关于房屋数目、建筑物类型、用途、人口等的信息，以便与数字化地图相联系，并随后划分人口普查点查区。这一活动的结果将是编制地理基准化的地图，用于人口普查。显示点查区位置以及建筑物、小巷、僻巷等的详细布局的地图将交到点查员手中，以进行点查。这些地图还将与其他政府机构分享，以执行它们各自的方案(如果了解更多信息，请联系Chinmoy Chakravorty: [cchakravorty.rgi@censusindia.gov.in](mailto:cchakravorty.rgi@censusindia.gov.in))。

## 第六章

### 发布人口普查结果、产品和服务的地理数据库

#### A. 引言

6.1. 第五章讨论了利用地理空间基础结构来支持人口普查的点查工作。第六章谈到点查工作后国家统计局办公室将承担的地理任务，以及发布和使用地理基准化的人口普查信息。在人口普查进程的这一点上，点查的结果应该已经出来了。在本《手册》中自始至终强调的一点是所有地理计划必须与人口普查的总体计划相一致。对国家统计局办公室投资于地理数据库来说，第二个关键问题是数据库能够被用于人口普查过程的所有阶段。一旦点查工作完成，可以利用点查结果来进一步改进数据库。利用地理数据库来制成产品，为公众提供信息和教育公众，这将引出集成和比例的问题，需要国家统计局办公室以外的机构做出贡献。

6.2. 如果有完善的数字化人口普查地理数据库，那么只需通过集成就可以为管理单位或统计单位建立统计数据库。对于在2010年人口普查不使用数字技术来制作点查区地图的国家来说，也有办法在这一阶段开发数字化的地理基准人口普查数据库，制作符合出版之要求的地图作为普查报告的配套图，送给需要从空间角度分析人口普查数据的外部用户使用或满足内部需要。数据库经过编制可满足各级管理层中某一适当层次的需求，或者满足其他大型统计地区的需要。在满足大型统计区的需要方面，制作一套数字化数据库所需要的资源比为点查区制作一个完整的数字化地图数据库要少。

6.3. 然而，第六章所讲的大多数内容都是建立在这样一个假设之上：为了做好人口普查中的点查工作，已经建立了一个完备的点查区或居住单元的数字化数据库。为了证实开发这样一个数据库进行投资确实有必要，国家统计局办公室需要用一种长远的眼光看问题。人口普查开始之后，立即需要进行的工作只不过是为了下次点查工作准备地理材料的第一步。

6.4. 第六章主要讨论的内容是，人口普查结束后以及两次人口普查之间的地理数据库的管理以及产品的开发与发布。其他涉及的主题包括：将实地发现的变化包括进来；将数据集成到发布单位的过程；数据库维护；发布产品和服务；披露和数据保密方面的考虑；营销；外联；地理产品；用于公布的人口普查地图和数据库；从空间角度分析人口普查数据的一些方法；以及与因特网绘图和人口普查数据库发布有关的问题。

## B. 点查结束后以及两次人口普查之间的工作

6.5. 点查之后是周密规划给予回报的时候。如果国家统计局办公室遵循了在这一过程的早期制定的全面计划并且坚持该计划，那么最终产品的确定引导了在点查阶段做出的选择。如果计划要求在点查区一级公布结果，那么可能不需要进行调整。对于在点查期间发现的地理变化需要另外进行一些调整。国家统计局办公室需要及早确定哪种错误需要立即引起关注，哪种错误可以先放一放，因为可能会延误结果的发布。

### 直接工作

#### 1. 整理点查员所做的更新和改动

6.6. 点查工作的一个方面就是培训人口普查员注意点查区划界中或者所查地域的底图特征中的错误或不一致之处。人口普查制图办公室应鼓励点查员记录错误，地方监督员可以在点查后收集这些错误，提交至人口普查制图办公室。人口普查地理单位随后可以通过国家统计局办公室设计的一个程序，根据该信息对用于点查区地图制作的地图数据库进行更正。这一程序有两个优点。

6.7. 首先，保证对数字化地图和硬拷贝地图产品进行制表和开发时，所使用的点查区地图轮廓与点查过程的实际情况一致。其次，把点查区边界的修改结果输入到总的地理数据库中，会对将来基于相同或类似的地理信息收集单位的人口普查工作或其他类似的数据收集工作提供方便。明显的优点是尽可能使地理数据库不要过时，同时也支持实地参与者改进整体数据质量的积极作用。不过，进行调整时不能不对整个人口普查计划所受到的影响做一些评估。

#### 2. 将收集单位与制表单位或统计单位进行整合

6.8. 点查工作结束后最重要的责任是，在根据人口普查结果开发表格统计数据方面提供支持。人口普查数据应能满足不同加总区的需求，因为不同部门的人口普查数据的使用者希望利用不同的加总区作为规划和运作的基础。因此，点查区应按照不同的报告单位加总在一起，根据需要开发出各种各样的人口普查产品。不可能所有用户都对覆盖全国的点查区级别数据库的详细程度有要求，更不必说文件管理的挑战性了。

6.9. 为了使数据收集单位(点查区)与制表单位在各个地理级别保持一致，需要开发等效或兼容文件。协调这两种不同的地理级别的表格有时也称为“整合表”。等效文件可为每个制表单位列出相应的点查区，点查区是产品输出单位的一部分。这种列表一旦确定，即可通过标准数据库操作进行加总。从一次人口普查到另一次人口普查的点查区或其他单位的改变见表VI. 1，从表中可以看出2000年的一个点查区(普查地段)在2010年被分成了三个新的点查区(人口普查地段)，编码随之做了修订。

表VI.1  
新旧点查区单位的对比

字段长度字段描述	
1-11	人口普查地段编码
12	区(P)旗
13-18	1990年人口普查地段的街侧英里数
19-22	1990年人口普查地段的街侧英里数的百分比
23-31	2000年人口普查地段编码
32-33	2000年人口普查地段后缀
34	2000年人口普查地区(P)旗
35-40	2000年人口普查地段的街侧英里数
41-44	2000年人口普查地段的街侧英里数的百分比
45-50	记录所覆盖的区域的街侧英里数
51-64	14 记录的土地面积(1 000平方米)
65-66	2 2000年州名缩写
67-126	60 2000年县名
下面的记录显示的是1990年的人口普查地段402在2000年被分成了三个人口普查地段：402.01、402.02和402.03	
10001040200	11735029610001040201 34771000 3477 9796
10001040200	11735044210001040202 51891000 5189 32059
10001040200	11735026210001040203 30691000 3069 59822

6. 10. 如果在处理过程中能采用一致的编码方案，开发等效文件的工作就会比较容易。本《手册》中，尤其是第三章中，详细介绍了人口普查单位的地理编码和属性问题。要在开展人口普查制图项目的早期阶段制定出一套直观和灵活的规定，为每一特定的点查区制定一个数字或字母数字，这一点十分重要。

6. 11. 需要开发等效文件的产品输出单位的数量可能非常多。除了法律和行政单位(例如地区或省外)，还需要为很多计划或运营单位进行普查数据的编译工作。例如保健部门、学校地区、运输规划区、选举区、公用设施区、邮政区以及环境规划部门等(见图VI. 1)。这些单位在有些情况下与行政区完全重合，但经常与标准报告单位不一致。此外，私营和学术部门有可能会提出一些特殊的制表要求。因此，开发一套用于生产和维护等效文件的统一程序，就成为了人口普查制图部门的一项重要任务。正如本手册通篇强调的，有效的规划对于国家统计局来说有许多可见和不可见的益处。

6. 12. 应该开发额外的兼容文件，目的是对过去和现在的点查或统计报告区进行协调。因为数据收集和制表单位经常会发生变化，人口普查数据的用户很难确定随着时间改变普查变量会有什么变化。因此，人口普查办公室的地理单位应当经常跟踪国家人口普查地理发生的改变，给数据用户提供兼容性文件，以协调过去和现在的人口普查数据。



图VI.1  
人口普查制表和报告单位的例子



6. 13. 在国家统计办公室参与国内的空间数据基础结构(SDI)活动时，这项工作特别有用，因为可以向那些需要人口普查数据的人提供所需的准确说明，以确定正确的人口普查文件的位置，用于他们的特殊目的。

### 3. 数据库的存档

6. 14. 当总地理数据库内的错误和不一致问题得到处理后，应制作所有地理数据集的基准拷贝并进行存档。为了反映人口普查时的情况，在普查地理方面进行冻结，由此产生的总数据库将作为所有地图输出的基础，包括基准地图、人口普查结果专题地图，以及用于发布传播目的从该总数据库中抽取出来的数据。点查后制成表格的所有人口普查结果将归入该数据库内的基准单元。这也意味着要对整个文档和元数据彻底检查一遍，以便将来人口普查办公室能够回答有关数据方面的任何问题。数据库工作一旦完成，应立即把基准数据库作备份并保存在一个安全的地方。

6. 15. 对于需要进行制图工作的国家统计办公室来说，这样的一套数据库可作为两次人口普查之间定期更新的基础。以下部分将对继续进行制图的各种好处进行讨论。

### 4. 数据库的维护：继续进行制图的好处

6. 16. 正如本《手册》前面指出的那样，只要能在人口普查核心任务之外多方面利用为普查目的建立的数据库，数字化地理普查方案的收益就会超过其成本，尽管成本是巨大的。只有维护数据库以便为未来人口普查应用进行的

更新只需要较少的资源，一系列的好处才能变成现实。只有国家人口普查制图方案具有高度的连续性，才能确保最大限度地发挥人口普查地图数据库的使用潜力，确保在以后的点查中尽可能使用现有数字化数据。因此，持续的人口普查地理活动可以确保数据库开发时的投资产生收益。

6. 17. 需要考虑的一个重要问题是，人口普查制图部门在普查完成时应立即进行数据库的维护工作。要持续不断地根据得到的新信息，对边界和其他地理特征进行更新。在两次人口普查的间隔期间，应该进行一次明确而系统的版本控制工作，记录并公布数据库中的改动情况。例如，只允许一个工作人员或一个工作人员小组对主数据库进行修改，避免不同人员对不同版本的数据库进行修改而造成需要调整的混乱局面。

6. 18. 在两次人口普查的间隔期间，人口普查制图机构应紧跟行业的发展趋势，了解其他人口普查机构采用的新方法。这有助于在知情的情况下对软硬件升级做出决策。鉴于技术的飞速发展，需要在这些领域不断投资，以确保间隔期的人口普查工作的高质量。

6. 19. 数字化制图数据的开发需要在计算机的使用、地理知识的掌握以及专业软件包的操作方面拥有特殊的专业技能。为使人员掌握最基本的地理概念和任务而提供培训的代价十分昂贵。因此，为了确保人口普查制图项目的长期成功，保持人员的连续性是一个关键性因素。人口普查办公室要确定一位骨干人员，负责人口普查间隔期间的数据库维护工作，为其他统计应用提供地理空间服务，如抽样调查，并充当机构存储器。通过这种方式，在下次点查工作中，实现人口普查地理信息的平稳运行。例如，可以让骨干人员对临时招聘的数字化处理人员或实地工作人员进行训练。保留骨干人员还能降低工作的启动成本，否则的话，需要招聘地理专业人员，而新招募的人员需要花费一定的时间才能完全投入到人口普查地图绘制工作中去(关于人力资源问题的详细考虑，见第二章)。

6. 20. 需要再次强调的是，对于人口普查制图工作有一种长期观点十分重要。从长期战略着眼，在两次人口普查的间隔期间，为了维持国家统计局办公室中的制图能力而补充资源是值得的。

## C. 发布地理人口普查产品

6. 21. 国家统计局办公室常常忽视发布工作，然而，这是教育公众和支持公众的活动的唯一最重要的途径。必须再次强调的是，国家统计局办公室应在人口普查过程的早期确定产品和服务系列，包括结构和格式。信息发布规划、披露和数据保密问题、营销、外联和教育等问题将在下文概述。

### 1. 规划数据的发布

6. 22. 确定地图产品和服务的输出以及发布时间需要与全部人口普查项目的时间安排密切协调。普查数据的制表需要来自人口普查地理单位提供的信

息，专题地图和数字化地理数据库只有在人口普查数据完成后才能实现。

6.23. 如何选择适合的产品应基于对客户需求的详细评估，也就是说要进行市场调查，这些工作应在普查规划的初期阶段完成。应尽早安排产品的发布计划并广泛公布，以获得用户的反馈意见。

6.24. 建立一个能对国家统计局办公室进行指导的、代表人口普查数据重要用户的顾问团十分有用，顾问团的作用并不仅限于人口普查规划阶段，而是可以形成一种正式或非正式的永久性机制，以利国家统计局办公室与数据用户交换意见。本《手册》的序言中提到了利用小区域人口普查数据的实例，从这些实例中可以看出广大数据用户的一些情况，人口普查办公室在其用户需求评估中应加以考虑。

6.25. 过去的经验证明，在确定输出产品和服务时，受人口普查数据用户欢迎仅能作为一种有局限性的参考。需求是变化的，部分原因在于数据用户的技术能力发生了变化。上一轮人口普查结束时还无法得到的数字化地图数据库产品和服务，而它们将是当前一轮普查最重要的的产品之一。目前很多国家对硬拷贝地图的需求仍较大，而对数字信息的需求较少，这种情况正在发生改变，因为地图产品的电子发布变得越来越普遍。人口普查制图机构将面对时刻变化的用户需求以及他们的特殊需要时，必须能够灵活地做出反应。

6.26. 在计划产品策略时，最好着眼于几年后的状况。例如，即使是在贫困国家，因特网也成为了一个重要的数据发布渠道，因为全世界的通信基础设施有了改善。随着新的数据产品的出现，新的用户群体也出现了。例如，移动电话公司寻求地理基准化的人口数据来规划新塔和其他基础设施的安装。为了增加人口普查数据采集带来的社会效益，人口普查机构要积极地发掘新的潜在用户群体，向他们介绍自己的产品。在政府和非政府数据用户中开展这一工作的一个有效途径是参与国家空间数据基础结构。

6.27. 人口普查办公室应评估对其产品和服务的可能的需求量，借以评估满足客户需求需要达到的能力。当然，这并不容易，因为随着新的产品和服务的引进，需求会增加，并且当新用户认识到这些产品和服务有可能满足其需求时，需求量还会增加。所以，人口普查办公室要做好准备，一旦提供产品和服务，要跟得上不断增长的需求。实际计划可以根据需求水平进行调整。建议对人口普查数据用户的需求尽早做出明确界定：哪些是**必须**满足的（即法律要求满足的），哪些是**应该**满足的（通过优质客户服务的概念），哪些是**不必**考虑的（因为资源限制）。清晰的优先顺序也有利于制定发布人口普查产品和服务的时间表。

6.28. 强烈建议国家统计局办公室通过开放的数据发布政策，尽可能广泛地提供结果、产品和服务。开放式的数据发布政策有助于减轻人口普查办公室的工作量。所谓开放式政策就是以低价格或免费提供数据访问。在有些国家，人口普查数据可以免费使用，私人服务提供商可以针对某些人口普查数据用户开发适合他们特殊需要的产品。这样，人口普查办公室就可以把精力放在法律法规要求他们为其提供服务的数据用户身上。

6. 29. 某些人口普查地理数据产品是为了满足内部和官方需要而制作的。其中可能包括等效文件和基准地图库，以及其他专项用途产品，如选举区地图等。在某些国家，可能要求人口普查办公室依据法律要求制作某些地图产品。这些产品可能必须定期制作，或根据特别要求制作，例如根据政府各部或议会的要求。

6. 30. 其他更为一般的产品被广泛地发布给政府部门或私人部门用户以及广大公众。人口普查办公室应尽量发掘发行渠道，通过出版分析报告以及数据和结果，满足用户的各种需求。下面提供了这类空间分析的一些例子。

6. 31. 以下将要探讨人口普查产品的输出类型以及如何发布的问题，包括按要求定制的产品、以硬拷贝或数字化形式发行的专题地图、数字化地图数据库的发布、数字化人口普查地图册以及基于因特网的信息发布策略。尽管很多此类产品的输出都要求具备一定的专题制图技巧，但本章以下的讨论内容仅涉及最一般的问题。附件四对专题地图的设计有更为全面的论述。

## 2. 披露和数据隐私性考虑：区差问题

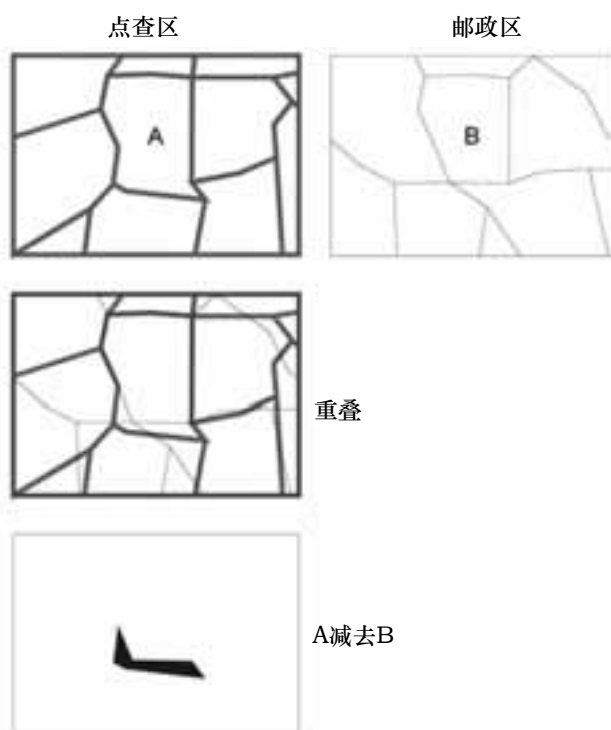
6. 32. 国家统计局办公室应该意识到公众对维护与人口普查有关的个人信息的隐私性的关注会影响到数据的发布。关于维护个人隐私性的准则包含在标准的联合国准则中(例如，见联合国，2008年)。

6. 33. 不同的政府机构和外部的数据用户可能需要将人口普查数据集按照不同的小地理单元进行划分。例如，有的机构采用小邮政地段或卫生保健区作为其基本的报告单位。为了满足这些数据用户的需要，国家普查机构要按照几个系列的小地理区域发布人口普查信息，而这些小的地理区域的边界并无联系。如果对两个或两个以上系列的区域发布边界和数据表，一个熟练的用户就可以利用地理空间操作和简单的数据表操作来得出极小的地理区域的人口普查统计数据。这些新单位的人口普查数据可能超出了机构的披露下限。这个问题被称为统计披露的“区差问题”(见Duke-Williams和Rees, 1998年)。

6. 34. 除非重叠地段之一为零，如果边界重叠得毫无规律，便不会发生这种问题。在大多数情况下，用户不敢肯定零值实际上是正确的。因为大多数人口普查办公室采用了干扰技术或宽带编码技术，防止用户推导出人口稀少地区的小群人口的具体特征数字(所谓宽带编码技术就是只给出一个数据范围，例如，“<10”，而不给出准确的小数字)。

6. 35. 如果来自某个地理小区地段的数据嵌入到另一个地理小区地段中，而且用户又有两组小组的数据表，那么就会出现区差问题。例如，在图VI.2中，邮政地段B嵌入到点查区A中。通过重叠两组边界，我们可以确定在A而不是B中的地理区域。通过数据表，只要从点查区A的计数中减掉邮政地段B的计数，就可推导出该剩余小区内的人口普查数据。即使A区和B区的数值均不超出泄密范围，新推导出来的小块地区则会明显超出。

图VI.2  
统计披露中的区差问题



6.36. 为了避免数据泄密，国家统计局办公室应仔细审查不同普查区域的边界划分。在可能出现区差现象的情况下，则要采取额外的数据保护措施。Duke-Williams和Rees (1998年) 仔细分析过区差问题。基于试验，他们提出了解决这个问题的一些普遍建议：

- 表格采用最小一级的临界值。有些进一步的防护措施是对非常小的区域引入大量的干扰数据，或者对于小的数量采用范围定义，而不使用精确的数值。这种方法将减少发布小于小区单位数据集的人口普查数据的危险。
- 对于选定发布的主要普查地理单位，应越通用越好。例如，如果国内的大多数机构使用小的行政管理区作为其主要基准，那么普查数据也应该按照这样的地理单位进行发布。
- 如果所发布的另一套地理区域远远大于主要的普查单位，风险则非常小。即使出现区差现象，最后的计数结果也不大可能超过安全范围。
- 如果分辨率大致相同的两个普查地理区域非常相似，也就是说大多数边界相同，出现区差风险的机会要大于边界非常不同的区域。
- 如果国家统计局办公室确定带已知坐标和边界线泛化水平的标准的行政或统计边界，区差问题则会减少。

### 3. 地理人口普查产品和服务的市场开发

6. 37. 由于对分类普查数据的需求越来越大，国家统计局办公室应考虑简化用户获得其产品和服务的途径。可以制定战略来发展内部能力，以满足用户需求或者利用外部的专业知识。国家统计局办公室应酌情开发内部能力以发布它自己的数据和分析，只有在潜在需求超过统计办公室的发布能力时才依靠外部实体，如商业公司。

6. 38. 在考虑是出售还是免费提供数据访问时，国家统计局办公室应权衡创收和扩大数据用户群这两个方面。有些国家打算收回人口普查地理数据库的开发成本，而国内对小区域统计数据也有较强的商业需求，因此它们需要发掘与私人数据销售商签订市场协议的可能性。潜在的合作者包括当地主要的地理信息系统软件生产厂家的发行商、在线制图和图像提供商以及地理人口数据销售商。大多数处于领先地位的地理信息系统销售商均制作和销售各种课题的地理数据集。这样不但能增加收入，同时也能促进人们使用它们的软件，其方法是用他们的软件数据格式提供数据集。这些私人销售商有时可与国家测绘和统计部门合作，共同制作专业规格的地理信息系统数据库。

6. 39. 对于国家统计局办公室来说，这样做有一定的好处。软件和数据销售商可以通过他们的技术专长和可能的计算资源促成开发地理数据库的发行软件，以换取数据库销售的收入份额。国际软件销售商可以参加，增加本国地理信息系统数据的销售。其他国家的需求可能来自国际公司或研究该国的科研机构。

6. 40. 在与商业软件销售商合作过程中，一个可能出现的问题是销售商在发布数据时只是用他们自己的专用格式。人口普查办公室应确保需要其他格式的数据用户同样能够访问这些数据。至于商业化发行的不合理之处前面已经探讨过。一旦把市场营销权授予了私人公司，国家统计局机构就无法以免费形式或低价格提供数据了。如果目的是尽量扩大传播，由部门本身自行开发与销售数据库则是最好的方法。

6. 41. 其他的潜在市场伙伴是大学或其他从事信息发布的政府部门。在任何情况下，必须签订明确的市场营销和收入分成协议，避免以后发生问题。人口普查办公室应根据数据的制作成本、广告宣传以及销售费用对数据的市场价值进行评估，确保达成一份既公平又能使双方共同受益的协议，这是公共部门与私营部门，或公共部门与另外的公共部门之间建立伙伴关系的基础。

### 4. 外联和教育

6. 42. 为了保证广泛了解数据可用性和尽可能广泛销售地理基准化普查数据，国家统计局办公室应制定一份对外宣传计划。该计划可以包括反映人口普查地图作用的打印的小册子和招贴画，广泛散发到中小学、大学、商业企业和国家及地方政府机构。可以在进行人口普查之前开展一个营销和发布活动，以提高对最新统计数据的认识并确保公众的全面参与。

人口普查办公室还可以组织一系列全国性的区域用户研讨会。在这些讲习班中，人口普查工作人员可以向广泛的潜在用户介绍用于分析人口普查数据的免费的或低成本的绘图软件包，同时告知公众人口普查工作将要产生的一系列产品。

## 5. 可能的产品清单

6. 43. 下面介绍了国家统计局办公室在完成点查工作后应发布的各种产品，包括等效和兼容文件、一个基准地图库、地名字典和矩心文件，以及主题地图。

### (a) 等效文件与兼容文件

6. 44. 等效文件前面已经提到过，是人口普查制图机构在点查工作结束后的首要职责之一。除了直接用于普查数据制表之外，等效文件也可作为一种产品输出。数据用户可能需要了解某些点查区是否属于特定统计区或行政管理区，或者哪些小区域统计单元能够组成更大的报告单位。

6. 45. 发行等效文件产品时，既要有硬拷贝也要有数字化形式的产品。大多数与数字化普查数据打交道的用户，无论是确定地理基准，还是查找数据表格，都可以从计算机可读格式中受益，在数据库操作中直接使用这些文件。

### (b) 基准地图库

6. 46. 除了等效文件外，普查机构还应当制作出所有报告单位的基准地图。在某些国家，人口普查制图部门依照立法规定，生产这样的地图，供政府官员和普通公众使用。

6. 47. 基准地图可以用数字化方式发行，以简单图形格式、外部页面语言或PDF格式发布。然而，并非所有的用户都能使用数字化文件。因此，也应根据要求以全套硬拷贝的形式提供基准地图。

6. 48. 基准地图需要配以文字说明，详细描述每一个普查地理区域的特征。综合性基准地图文件的一个良好范例是美国人口普查局出版的《地理区域基准手册》(可上网查阅：[www.census.gov/geo/www/garm.html](http://www.census.gov/geo/www/garm.html))。

### (c) 地名字典和矩心文件

6. 49. 通常应由国家制图机构负责发行地名字典，上面列出各地地名及其地理位置。然而，以人口普查为目的开展的大型国家地图绘制项目将为国家地名字典提供一个改进或信息更新的基础。在某些国家，这类数据没有其他来源可供使用，只能从人口普查地图绘制项目获得地名字典。如果人口普查地图绘制项目已经进行过广泛的全球定位系统数据的收集工作，那么开发出一部列出所有地理位置的地名字典应当是顺理成章的事。一部地名字典可以包括所有居民区的名称以及别名、经纬度、行政编码和人口。

6. 50. 一部地名字典应以数字化形式存储和传播，允许在地理信息系统中直接使用地理坐标和地名信息。开发一个简单查询系统非常有用，用户可以找到一个特定地点的坐标，例如，查找一个特定省份内的某一个村庄的坐标。这类数据可以通过因特网提供，只要有一个简单的前端连接数据库即可实现。数字化边界文件也是一种可能的普查产品。下面将讨论这些产品。

## 6. 用于出版目的的专题地图

### (a) 地图的威力

6. 51. 专题地图可以作为人口普查出版物印制。在讨论各种类型的专题地图之前，有必要指出为什么专题地图对于展示人口普查结果十分有效，原因如下：

- 地图能够传达一种概念或理念。
- 地图经常意味着对文本信息提供支持，有些难以用文字解释的专题或问题用地图或图形就可以更好地说明。
- 地图能引起看图者的好奇心。在报告中，地图能起到抓住人们视觉的作用，引起读者的注意，鼓励他们阅读与之配合的文字。
- 地图的简洁形式中囊括了大量信息。地图不仅能表达巨大的数字信息，而且还能说明观察物质间的空间关系。其他方式很难与地图的这种能力相比。例如，一幅表示中国或美国人口密度的地图，能显示3 000多个数据，将这样一幅地图打印在信纸大小的纸上，基本上不会对清晰度造成损失。但要把3 000多个数据塞进一页信纸，不但不容易，而且还会丢失信息。比如，高密度和低密度人口区分别在何处这个问题，数字难以直接表现，而地图却一目了然。
- 地图可以用于描述、探索、确认、制表，甚至装饰目的。地图可以服务于多种意向。在人口普查报告中加上地图可以起到说明的作用。这些地图只是单纯罗列普查结果，有时附加，有时不附加任何说明。普查数据在人口学家或地理学家的手中时可能会得到较为深入的运用，他们或许会通过地图查找不同变量之间的关系，比如，预期寿命与识字率之间有什么关系等。在最终的报告中，除了文字说明和图表之外，含有此类因素的地图可以派上用场，用来支助分析人员的结果。这样，地图可以成为一种工具，用以确认某种结果。当然，仅仅依靠观察地图不一定能得出有意义的结论。地图可以简单用于汇编目的，例如，用于显示一个国家所有学校或诊所的数量。当然，有了汇编目录，便可很快做出分析，例如，可以从中发现哪些地区的公共设施不足，等等。最后，地图还是一种时尚用品，因为它们通常很美观。我们可以看到大批地图悬挂在办公室的墙上，但几乎看不到把统计图表或数字表格挂在墙上的情况。



- 地图使人们愿意做出比较。无论是描述性还是探查性的专题地图，其重要用途就是跨越地理空间对事物进行比较。能够比较的事情很多，如：
  - 同一地图上的不同区域之间：哪里人口密度最高？
  - 在不同的地图之间：甲省的儿童死亡率是否高于乙省？
  - 对于同一地区的不同变量之间：不同地区的男女识字率是否存在差别，差别出现在哪一地区，差别有多大？
  - 在不同时期的地图之间：自上次人口普查以来，出生率是否出现下降？

(b) 人口普查数据专题地图

6. 52. 地理空间技术在人口普查过程中的使用促使人们形成一种看地图的新观念，它与传统的地图学具有很大差异。在一台计算机上，可以很快生成多种地图并将其显示在屏幕上。它支持新的工作模式，对数据的有效性、数据格式的探索以及数据分析均可进行优化，这个领域越来越多地被称为“地学可视化”。地学可视化源于地图制作和显示的既定原则。它被定义为以地理测量尺度，创造和使用直观表示法来方便关于人为环境和自然环境的思考、理解和知识构建(Longley及其他人，2005年)。

6. 53. 建立在计算机屏幕上的地图，通常被称为“虚拟地图”，以区别于印制或手绘的硬拷贝地图，尽管计算机地图和硬拷贝地图之间的界限正日益模糊。在人口普查过程中，在数字化普查地图绘制的早期阶段，不必过多地关注传统地图的绘制设计。正如第三章所说的那样，要把工作重点放在数据库的开发和验证方面。即使是点查区地图的制作，通常也只是做些相对简单的绘图设计工作。点查区地图表示出一个点查员工作区的主要特征。

6. 54. 一旦人口普查数据经过编制，人口普查办公室一般就要制作出符合出版质量要求的地图，用以展示普查结果，同时发布普查报告。这样的地图将面向范围更广泛的非专业群体。因此，要对地图进行更仔细的加工和设计，最终产品可以印制成册，也可以做成只读光盘或DVD，或在因特网网站上发布。

6. 55. 表VI. 2列出了一些可供选择的专题地图类型，可将它们放入人口普查地图册或放到人口普查机构的网站上。很多其他类型的地图也可以考虑按专题出版，或者突出普查结果所涉及的国家地区有吸引力的方面。正如人口普查数据的分类制表那样，可根据性别、年龄或城镇/农村等进行分类，也可将地图分为各种不同的人口组分。如果以前人口普查的可比指标仍然有效，将不同时期的地图加以比较也能获得大量信息(专题制图原则的概述见附件五)。

表VI.2

## 可用于人口普查地图册的专题地图清单

专题地图的类别/名称
人口动态与分布
不同人口普查期之间人口变化的百分比
平均年增长率
人口密度(每平方公里人口数)
城市人口占总人口的百分比
主要城镇的大小及分布
迁入率、迁出率和净迁移率
本国出生和国外出生数量
本国其他地区的出生数量
人口学特征
性别比率(每100个女性所对应的男性人数),也可以按年龄组
年龄在0至14岁的人口百分比
年龄在15至64岁的人口百分比
年龄在65岁及超过65岁的人口百分比
育龄(15至49岁)女性的人口百分比
总扶养比率(年龄在0至14岁的人口与年龄在65岁及超过65岁的人口之和与年龄在15至64岁的人口数的百分比)
婚姻状况
出生率
总生育率
平均初婚年龄
死亡率
婴儿死亡率
自出生计算的预期寿命
残疾人口的百分比
社会经济特征
未上小学的儿童百分比
成人识字者的比率(15岁及15岁以上)
学校教育的平均年数(25岁及25岁以上)
15岁及15岁以上人口的文盲比率
15岁及15岁以上文盲人口
10岁及10岁以上人口的教育水平
劳动力占总人口的百分比
妇女占成年劳动力的比例
按经济部门、工作种类和就业状况列示的劳动力百分比
家庭与住房
每个家庭平均人数
户主为女性的家庭比例
每个家庭的居室平均数量
住宅占有状况(自有、租赁,等等)
建筑材料类型
能够获得适当住所的人口比例
能够使用洁净水的人口比例
能够使用电的人口比例
能够利用卫生设施的人口比例
能够获得保健服务的人口比例

资料来源:联合国(2008年)。

6. 56. 符合出版质量的人口普查地图通常只为那些相对集中的统计报告单位印制。人口普查部门可以出版国家的总体地图，根据省或地区说明各种指标的分布。此外，也可出版具体省份的详细地图。对于主要的城市地区，可以利用人口普查点或点查区一级的数据制作出非常详尽的地图。

6. 57. 地理信息系统和桌面制图软件包提供各种各样的制图功能，很多商业地图制作商已经转向全数字化的制作技术。然而，为了生产出高质量的地图产品，仍然需要相当丰富的经验和专业技能。计算机制图系统提供的工具并不能替代制图方面的训练。由于可以利用便于使用的制图软件包，导致地图大量增加，从而破坏了很多标准制图的设计原则。起初，人们把这种现象归结为早期地理信息系统软件包缺乏合适的制图功能。但尽管地理信息系统的程序越来越强大，功能也越来越多，没有经过或几乎没有经过制图技术训练的人也很难制作出像样的地图。

6. 58. 在大多数人口普查机构中，由专业制图人员负责地图的制作、出版和发行。在接受了一些数字化制图技术的培训之后，他们可以在计算机上制作高质量的地图。

6. 59. 由于地理信息系统和桌面制图软件的广泛传播，日益增多的专题地图是由那些在地图设计原理方面未经过训练或几乎未受过训练的专门人员制做出来的。附件五中提供的信息不论是对于负责制图的骨干人员，还是那些在人口普查机构内外通过数字化空间数据库偶尔进行一下制图工作的人都是有益的。制图学和专题制图方面的出色参考书包括以下这些人的著作：Robinson和其他人(1995年)、Kraak和Ormeling(1997年)以及Dent(1999年)。MacEachren(1994年)曾专门针对在制图学方面缺少正规训练的地理信息系统用户，出版了一本关于专题制图方法的有用的入门书。专门为地理信息系统设计的设计指南正在变得越来越普遍(例如，见Krygier和Wood, 2005年；以及Brewer, 2005年)。

#### (c) 数字化人口普查地图册

6. 60. 通用性更强的地理数据库的使用者是具有相当的地理信息系统操作经验的用户，而数字化的人口普查地图册则面向一般公众、学校以及其他非专业用户。下面介绍两种制作数字化人口普查地图册的方法。静态的人口普查地图册包括一组地图和人口普查部门编写的其他材料。其本质是用户可以改变查看呈现顺序，但无法改变呈现的内容本身。而动态的人口普查地图册把数字化地理数据库与人口普查数据结合在一起，构成一个简单的制图软件包。用户可以利用这些数据制作出用户需要的地图，这些地图可以打印输出或复制到其他的应用软件中。

##### (i) 静态地图册

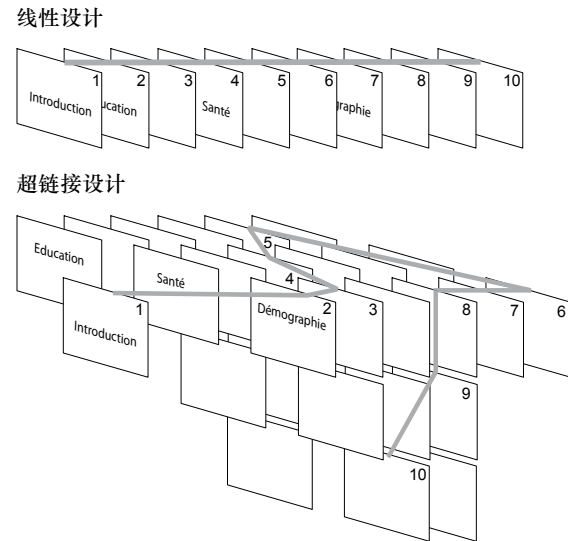
6. 61. 一个静态的数字化人口普查地图册可以把地图、表格、图表以及照片或电影剪辑等多媒体产品集成到一起，形成漂亮而方便的使用环境。可以把这些内容放入一个标准的展示软件中，如PowerPoint幻灯片。有些图形

展示软件允许开发者制成一套可单独运行的图形展示产品，与免费的查看软件一起发行。大多数的展示画面或图形均能导出，制成PDF格式的文件，通过计算机可读媒介或因特网进行传播。地图由桌面地图绘制软件完成，借助一种图形交换格式，或在Windows环境下直接通过剪贴命令将其送到展示软件中。

6. 62. 另一种展示平台是因特网浏览器。大多数计算机用户的机器里都有因特网浏览器，既可以利用它查看本地计算机上的文件，也可查看远程内容。地图和其他图形内容可以包括在GIF或JPEG格式的图片里，通过地理信息系统的地图版面生成。经过设计编排可以构成线性系列展示，引导用户查看按照一条主线安排的一系列地图和图形。这种形式可用于相对较短的展示。如果展示较大数量的地图时也使用这种形式，人们必须有足够的耐心，等待那些他们不感兴趣的幻灯片放映完毕。

6. 63. 大多数展示软件均可提供一种更好的基于超链接的设计选择。这种链接方式允许用户跨越不同的展示部分，还允许集成可能只令少数用户感兴趣的额外资源和信息。例如，在一幅有关地区人口预测的地图上，可以添加对说明预测假设条件的方法页的链接。

图VI.3 静态数字化人口普查地图册的不同设计展示方法



6. 64. 超链接的概念如图VI. 3所示，图中与线性设计进行了对比。在超链接设计中，列出若干平行的主题，根据具体需要利用链接把它们联系在一起。例如，排在简介页(1)之后的是三个平行的主线或章节：教育、健康和人口指标。用户可沿着一条路线(图中用粗线条表示)从人口统计(2)开始，其中一张幻灯片(3)显示15岁以下人口比例的地图、表格和图形。从这里，可以链接到显示儿童健康指标(4)、教育设施(5)等的地图。

6. 65. 超链接设计要求对展示内容进行非常仔细的设计，因为在相当多的链接之后很容易迷失方向。因此，在每一页上设计一个明确的导航工具十分重要，包括“面包屑”，设计这样的链接是为了让用户退出当前页面，找到自己的路径。

6. 66. 任何使用过因特网的人对于超链接都非常熟悉。一个静态的人口普查地图册也可以借助标准因特网浏览器的超文本语言予以实现，不一定非要使用展示软件包。网页设计工具可以使开发人员在设计人口普查数据库时具有相当大的灵活性。有一种工具可以使展示变得趣味盎然，例如，制作一幅可点击的地图。开始进入时屏幕上显示的是全国地图，根据指示点击有兴趣的省份，则可获得更详细的下一级地图。网络技术也允许包括多媒体的内容，也可以链接当前展示以外的信息，例如，连接到人口普查办公室网页的其他部分，或者连接到其他的政府部门。当然，用户只有借助因特网才能对这些内容进行访问。

6. 67. 使用网络设计工具有一个优越之处，就是静态的人口普查地图册可以独立运行，既可放入只读光盘或磁盘上发行，也可以放到人口普查部门的网站上供世界各地的用户浏览。

#### (ii) 动态地图册

6. 68. 除了静态的人口普查地图册之外，另一个选择是与制图软件一起发行数字化地图和数据库，让用户自己设计制作人口普查指标地图。当然，这种方法要求用户具有一定的地图绘制知识。一个动态的人口普查地图册包括用于迅速制图和减少磁盘空间占用的数字化边界文件，它的分辨率低于完整的人口普查数据库的分辨率。这种封闭集成式的属性表只包含一些选定的人口普查指标。用于制图的密度和比率应已计算完毕。

6. 69. 这种方法适合于那些不具备利用完整的数字化人口普查地理空间数据库所需要的地理信息系统专门知识和技能、但又想在探索和利用地理人口普查信息方面更具灵活性的用户，预先打包的静态的人口普查地图册不能提供这样的灵活性。

6. 70. 当然，问题是这样的用户可能得不到地理信息系统桌面制图软件包。因此，数据提供商应该随边界文件和数据一起提供一套易于使用的软件。易于使用的意思是只需要很少的培训和经验即可使用。说得更具体一些，这种应用软件应是“即插即用”型的，安装之后，用户立即就能用它制作地图。

6. 71. 有些人口普查办公室自行开发地图察看软件，并将软件与他们的普查数据产品一起发行。然而，维护这样一个软件既十分昂贵又占用资源，不能将资源全部用于数据的开发与发布。有些地理信息系统销售商目前正在出售地理信息系统软件工具箱，把它们放在一起就可以生成用户自行开发的应用软件，或者将地理信息系统功能集成在其他软件产品之中(例如，电子表格或数据库应用软件等)。

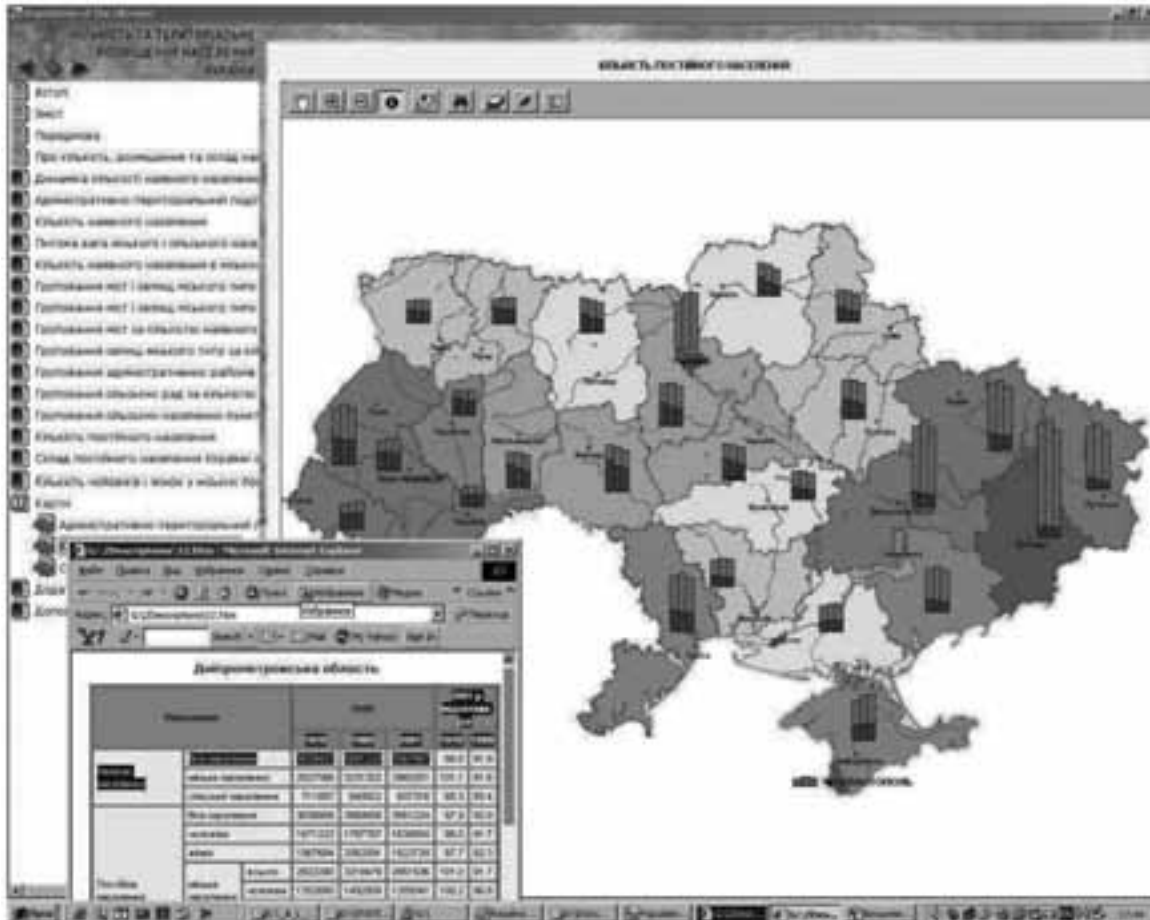
6. 72. 现在还有几种制图软件可供使用，它们是免费的并且和数据库一起发行(详情见第三章)。

6. 73. 某些商业地理信息系统销售商也制造免费使用的查看软件，并且允许用户在一个数据库发行软件包中免费发行这些简单的制图系统。实例之一就是由美国加利福尼亚Redlands的ESRI生产的ArcGIS Explorer软件包。ArcGIS Explorer是一个制图界面，处理由ArcGIS软件包生成的数据。

6. 74. ArcGIS Explorer界面很容易使用，系统提供了基本的制图功能，可用于制作专题地图，制做出来的地图可作为位图或Windows元文件导出。ArcGIS Explorer可以从本地硬盘或只读光盘读取数据。如果将计算机连接到因特网，该软件也可以显示远程网站上的数据。软件的分析功能虽然有限，但支持不同类型的数据查询，包括交互式或SQL指令查询，此外还支持地址匹配。

6. 75. 动态人口普查地图册的文档需包括与综合性人口普查地理信息系统数据库几乎相同的信息。但要记住，这些文档是为非专业用户编制的。一定要避免使用地理信息系统的专门术语。因为，这些用户不太可能把数据库用于

图VI.4  
乌克兰动态地图集中的一个屏幕截图



更高级的目的，所以这些文档的重点要更多地放在属性信息上，而不是过多地涉及地理技术细节。

6. 76. 乌克兰国家统计局委员会制作的图VI. 4说明了统计和图形要素在相同的地图展示中的使用。

#### (d) 空间分析技术

6. 77. 尽管空间分析有时在点查阶段使用(例如，聚类可以帮助查明要细查的居住单位)，但空间分析主要用于人口普查产品和服务。各种技术，包括缓冲、直线内插法、点模式分析以及统计地图，提供了超出标准专题(等值线)地图绘制的功能，现在在商业软件和开放源软件中都可以利用许多工具。

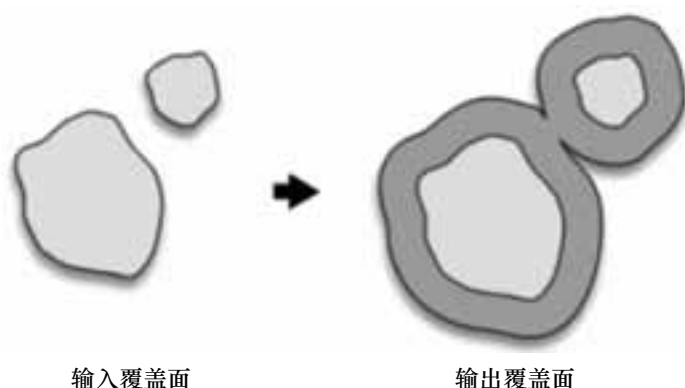
6. 78. 使用空间分析新方法背后的假设是，可以在点查区、人口聚居区或其他小区域级别得到“粒度”(或空间特性)级别高于从前的人口数据。如果分析人员或其他地理信息系统用户希望分析人口的空间分布或地图的人口变量或其他变量与别的变量的关系，他们如今可以利用各种技术，包括简单查询、对变化的衡量、描述性摘要和模型。

6. 79. Longley和其他人(2005年)把空间分析定义为一系列方法，当分析对象的位置变化时，使用这些方法得到的结果也会改变。空间分析有时被称为“地理信息系统的症结”，因为它包括将数据转变为信息的方法。可以利用空间分析来回答比如下面这些问题：X与Y的空间关系是什么？在哪里可以找到类似的特征？可以找到一个模型来说明一种普遍模式吗？不规则之处在哪里？

6. 80. 一些利用人口数据的特别有用的常用空间分析形式包括：

- “查询”被认为是最基本的分析操作类型。查询利用地理信息系统程序来回答用户提出的简单问题，不改变数据库，也不生成新数据。查询常常是分析的第一步，在查询中，人们试图建立一个带某些特性的单位子集，例如居民区，让用户可以检查一个观察对象与其他观察对象相比有多典型。利用进行了地理编码的人口普查数据进行的一个查询例子是：“选出人口数大于1 000的所有城镇”。随后，例如，可以概括这些城镇的属性，以衡量它们的总生育率与较小城镇和乡村的对比，随后可以对结果制图。“探索性数据分析”一词指的是利用诸如查询之类的技术来研究数据模式和趋势。
- 度量分析使用人口数据的位置特性，包括对象的简单特性，例如长度、面积或形状以及成对对象之间的关系，例如距离或方向，来描述该数据的各个方面。利用城市、城镇和乡村的矩心(或中心点)，很容易借助所有地理信息系统程序进行“距离度量”。可以挑选与一个学校、诊所或水源的距离超过1公里的乡村进行分析。然后利用居民区本身的属性信息，可以做进一步分析。

图VI.5  
一个多边形对象的缓冲



6. 81. 在前两个例子中，对数据集进行了查询或衡量，但没有产生新数据。转化是利用简单的几何、算术或逻辑规则建立新的数据集的空间分析方法。转化可以包括将光栅转变为向量数据，或将一系列地理信息系统坐标转换为一条路线或边界的操作。

6. 82. 在所有转换技术中，“缓冲”是最知名的，也是最重要的。缓冲包括通过确定距原物一定距离内所有区域来建立一个新的数据层。可以在点、线和多边形上进行缓冲操作(见图VI.5)，可以通过属性值来估量缓冲。例如，可以利用缓冲来建立旅行时间的模型，即围绕某一特征，例如一所学校或一个诊所，建立一个“集水区”。这样可以衡量可达到性，可在一国范围内对此制作地图。不过，可以联系其他数据集来使用人口数据，例如显示有可能发洪水的地域的数据集。由此进行的分析随后可以确定面临风险的人口和定居点，这些人口和定居点可能成为减灾方案的考虑对象。

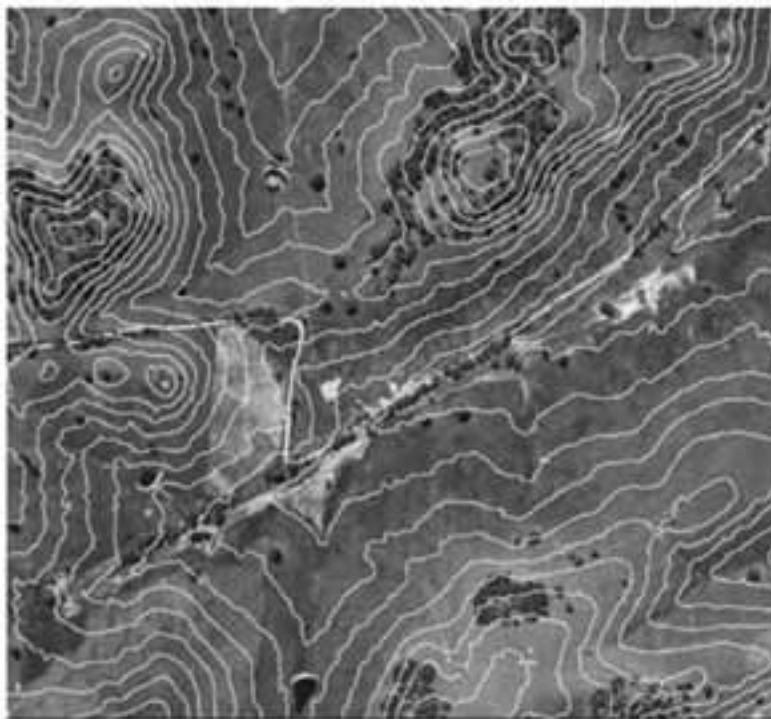
6. 83. 另一个转化例子是“多边形中的点”分析，就是确定一个点是在多边形的里边还是外边。可以利用这样的分析来比较位于热带风暴途经地区或地震区等危险地区里面和外面的进行了地理编码的村庄重心。“多边形重叠”分析包括对比两个不同的多边形数据层的位置。例如，两个行政区的边界可进行对比，以查明实地查过程中的错误。

6. 84. “空间插值”是一种空间分析方法，用于填补观察值之间的数值。各种方法，包括反距离加权法和“克里金法”被用来估计未取样的地点的数值，依据的是Tobler第一原理，即所有近处的物体都比远处的物体更相似。在克里金方法中，对一个表面的一般特性建立模型，以估计该表面的缺失部分(关于通过直线内插法根据一个卫星图像得出轮廓线的说明，见图VI.6)。

6. 85. 泰森多边形是空间物体，用于根据二维空间中排列的点之间的距离，围绕点数据创建区域。这一方法假设未取样的数据的值与取样的点的值相等(关于泰森多边形的说明，见图VI.7)。

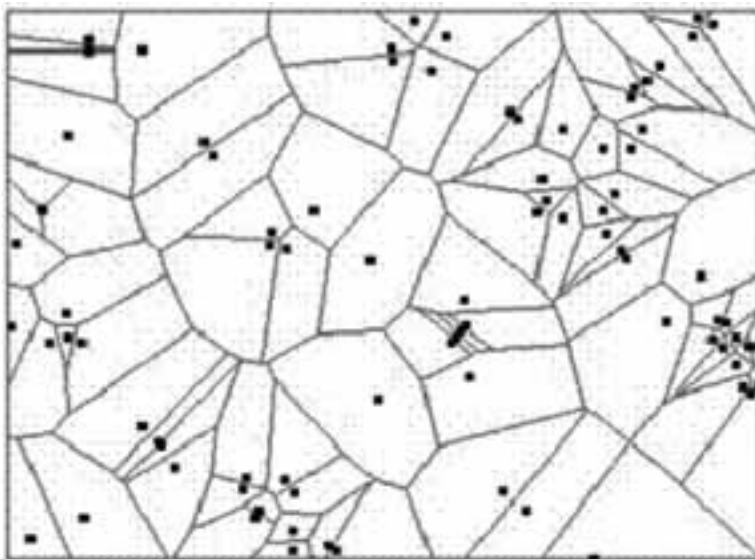


图VI.6  
创建轮廓线的直线内插法例子



6. 86. 描述性摘要描述性统计的空间相等物(例如平均数和标准偏差), 代表一个以 1 或 2 表示的数据集的实质。“人口中心”是一个统计平均数的二维相等物, 常常用来显示人口的中心, 利用的是居民点的 X 和 Y 坐标的加权平均数。“点模式”或“聚类分析”关心的是点在空间的分布, 无论其实际位置如何, 目的是确定模式是随机的、聚类的还是分散的。Moran 统计量 (0 和 1 之

图VI.7  
关于泰森多边形的说明



间)表示属性模式的一般特性, 可用来确定“热点”, 即高值围绕高值的地方, 或者冰点, 即低值围绕低值的地方。这些对于确定高危人群同样特别有用。

6. 87. “统计地图”(见图VI. 8)有时用来显示人口普查结果。在统计地图中, 原始多边形的面积根据其属性值, 例如人口规模或投票习惯, 进行了扩大或压缩。分析人员如今可以从ESRI和其他地方获得脚本和扩展(例如, 见MAPresso: [www.mapresso.com](http://www.mapresso.com)), 利用自己的多边形数据层创建统计地图。

图VI.8  
统计地图的例子



(c) 地图制作和公布问题: 输出的类型

6. 88. 在人口普查完成之后, 统计部门将出品一批用于各种目的和符合出版质量要求的地图产品。例如:

- 标准基准地图, 图中对人口普查数据汇总期间规定的每一个统计发布单位进行说明。
- 插图地图, 即在有关人口普查结果和普查方法的印刷报告中作为插图的地图。在这种情况下, 地图并非该出版物中的主要内容, 而只是对文本的补充。通常采用黑白图, 因为黑白图比全彩图容易制作。为了广泛地传播, 此类印刷品的印制数量相对较大, 一般由人口普查部门的印刷车间印刷或包给外面的印刷厂印刷。
- 印刷的人口普查地图册, 从小册子式的出版物到包括数十张甚至数百张地图的综合的硬拷贝地图册不等。
- 数字化人口普查地图册。如上所述, 这是一种成本效益很高的出版物, 在广泛使用计算机的国家可以替代印刷品。此类普查地图册可以

基于预先准备的静态地图，或者一个简单的专题地图制作界面，用户可在该界面上选择地图的各种变量、分类方案、绘图符号和颜色以及基本版面布置。

- 大多数地图可在因特网上发布，也可以印刷形式发布。要在静态地图和动态图中做出选择。静态地图与因特网上发布的其他图像或照片没有什么区别，而动态绘图软件则把专题地图的设计过程交给用户。
- 各种格式的特殊用途地图，这是为了满足机构内或机构外的人口普查数据用户提出的要求而专门制作的。这种印刷品的数量不多，均通过本单位内的设备制作，例如激光或喷墨打印机。
- 介绍材料。例如有关人口普查专题的幻灯片或大型招贴画，这类材料多得益于其中包含的地图内容。

#### (f) 绘图工具和软件

6.89. 新的先进的地理信息系统软件使得制成高质量地图的工作容易多了。不过，还是要使用制图人员传统上使用的相同的设计原理来制作有视觉效果地图。

6.90. 第一代地理信息系统软件包并未提供方便的地图绘制工具。一般通过命令行式界面或宏语言进行地图绘制。为了把文本加到地图上，用户必须在地图页面上为文本的所占位置规定坐标值，文本字号和式样需要使用单独的指令。新一代的桌面制图软件包在制图设计功能方面有了很大改进。用户在地图绘制过程中可以选用多种字体、线条、各种填充图案和剪贴画。系统本身还配备了地图专用符号集，包括地形图和专题地图常用的点线符号。桌面制图软件的用户界面与标准图形软件的界面基本相同，用户可以从交互式菜单中选择不同的式样，通过鼠标对地图上的各种元素进行移动和缩放。屏幕显示与纸张上打印出的地图看上去相当一致。

6.91. 目前的桌面制图系统和地理信息系统软件包所具有的地图设计功能可以满足大多数用户的需求。然而，在某些应用中，专业的制图人员仍然喜欢把底图从地理信息系统中导出，然后将其导入到图形设计系统、桌面出版软件或图形软件包中。这类软件包提供先进的图形功能，例如晕染地形描述、渐变填充或透明化处理，这些功能为绘图人员的设计带来了更大的灵活性。从地理信息系统向图形软件移动数据有两种方法可供选择。一是通过Windows操作环境所提供的标准剪贴功能，二是借助一种可以从图形软件导入的标准格式的中间文件。下面是对输出方式的介绍。

### 7. 输出方式：数字化文件

6.92. 下文将讨论数字文件的数据发布的各种格式。分析人员应该了解图形格式(符合出版质量的最终格式)与数据格式(形式上更粗糙)的发布有何不同。

6. 93. 所有地理信息系统和图形软件包都有让用户把地图版面导出为不同图形文件格式的功能。出于很多原因,这种方式十分有用,因为它可以使不同软件包相互交换文件。例如,地理信息系统的底图和统计软件包的数字表格均可导入到同一个图形软件包中,并在其中设计出最终的页面版式。文本处理软件可以导入成品图形,并将其集成到报告或出版物中。本《手册》中提供的大多数图形都是通过这种方法制作而成的。图形文件可以上传到网站上,作为静态地图图片,此外,还可作为文件附录通过电子邮件发送。

6. 94. 图形文件格式(类似于地理信息系统中的数据结构)可以分为不同种类,有的支持向量图形,有的支持光栅图像或图像文件。分布在规则网格上的细点或像素构成色彩及灰度上各不相同的光栅图像。连续变化的彩色色调或者灰度等级可以构成一幅照片式的图像。专题地图一般使用离散对象,因此显示时只需使用不多的几种颜色。

6. 95. 向量图形格式以点、线、面通过内部坐标系表现图形对象,内部坐标系或受输出的幅面限制,或不受设备制约。有些文件格式可以同时处理光栅图像和向量对象。这些格式对于地理信息系统绘图十分有用,例如,可将卫星图像与线条和多边形数据层相结合。无论是光栅还是向量图形格式,在将信息显示在屏幕上或输出到打印机之前,需要对图形内容进行光栅化处理,屏幕和打印机都是重要的光栅显示设备。计算机的操作系统和打印机的驱动程序会自动完成这一切。

6. 96. 以下将要对常见的文件格式做一个简单说明。该清单并不详尽,因为现行的文件格式不下几十种。

#### (a) 数据格式

6. 97. 向量文件格式与地理信息系统内向量数据的关系更为密切。这种格式可以更为紧凑地表现线条或多边形数据,并能充分地保留原始的地理信息系统数据层的分辨率。一些标准向量图形格式如下:

- WMF视窗元文件格式。Windows元文件是一种用于Windows操作系统的图形文件格式,经常用于向量数据,但也能存储位图图像。增强型的WMF(EMF)文件是一种更全面的WMF格式,开发这种格式主要是为了用于32位的Windows操作环境。在Windows的应用软件中,WMF是一种最稳定的文件格式,经常用于导入和导出图形文件。WMF格式还是Windows经常使用的一种图形格式,可以通过它把图形对象复制到剪贴板上,然后粘贴到其他应用文件中。
- VML。Vector Markup Language(VML)是一种XML语言,用来制作向量图型。VML是微软、Macromedia等在1998年作为拟议标准向W3C提出的,但作为一种网络标准遭到了否决,因为Adobe、Sun等提出与之相竞争的提议,即PGML。两种标准被结合在一起,创建了SVG。尽管作为一种标准遭到了W3C的否决,并且不为大多数开发者所了解,微软还是在Internet Explorer 5.0及更高版本和Microsoft Office 2000及更

高版本中执行VML。在Internet Explorer 5上运行时，Google Maps目前使用VML来表现向量。

- **CGM**计算机图形元文件格式。计算机图形元文件是一种用于存储二维图形数据的国际标准格式。当初作为一种纯向量标准开发，但后来的版本也支持光栅图像。有三种类型的CGM格式：一种是字符编码器，以减少文件大小和提高传输速度；一种是二进制制，用于提高存取速度；第三种是明码通信文本模式，用于基于文件的编辑操作。
- **HPGL**惠普公司图形语言。惠普公司图形语言是一种文件格式，当初是用于笔式绘图仪。在大版面喷墨打印机和静电打印机出现以前，笔式绘图仪是地理信息系统项目在打印大版面地图时广泛使用的一种输出装置。
- **DXF**绘图交换格式。绘图交换格式是由Autodesk公司开发出的一种文件格式。该公司是一家软件开发商，擅长计算机辅助设计和地理信息系统软件。当初设计这种格式是为了能让Autodesk自己的文件进行跨平台交换。现在，DXF已经成为一种标准的交换格式，受到大多数地理信息系统软件包和很多图形软件包的支持。
- **PS**外部页面格式和**EPS**增强型外部页面格式。外部页面格式实际上是一种编程语言，用于描述纯文本文件的向量数据。它是使用最广泛的页面设置说明。外部页面格式由Adobe图形软件公司开发。对与大小无关的向量图形优化之后，外部页面格式文件也可用来处理光栅图像。这种格式的主要用途是作为输出格式向外部页面格式的打印机发送文件和图形。因此，外部页面格式基本上是一种输出格式。很多图形软件包也支持外部页面格式的导入，但由于外部页面格式的编码没有达到完全标准化，所以这样导入的外部页面格式如果是在不同的计算机程序中创建的，则一般不能再做进一步的编辑。当跨越硬件平台传递外部页面格式的文件时，更会出现这种情况。有时，甚至在同一个软件中创建的外部页面格式文件也无法导入。

虽然常常无法对导入的外部页面格式文件做进一步的修改，但大多数的软件包都能将外部页面格式文件加入到一个文件中。但屏幕显示时看不到外部页面格式文件的内容，只能显示一个带有标记的方框。一旦将该文件送到一台支持外部页面格式的打印机上时，实际存在的外部页面格式文件的内容便会打印出来。由于外部页面格式文件与尺寸大小无关，所以导入的外部页面格式图形可以在缩放之后，充填到合适的空间内(关于PDF(可移植文档格式)，见下文第6.102段)。

#### (b) 光栅图像格式

6.98. 地理信息系统或图形软件包均可以直接建立光栅图像。在有些情况下，另外有两种方法对建立图像非常有用。一种是在支持光栅图像的软件中使用屏幕抓取命令。这类“屏幕抓图程序”有时在保留原始色彩方面比地理信息

系统或图形软件包内的导出功能还要优越。另一种选择是通过专业软件或硬件把图形对象转换成光栅图像。此类光栅图像处理器可以生成分辨率非常高的图像，能够保留向量格式图像的所有细节。然而，最后形成的输出文件可能非常大。

6. 99. 文件的大小取决于两个因素：图像中的颜色数量和图像的压缩程度。例如，某一个图像格式只支持两种颜色（黑与白），那么，每个像素只需一位便可表达。用8位（一字节）表达一个像素时，则可以存储256种颜色。高端显示器或高级图像格式用24或32位代表一个像素，则可以表现1 600万种颜色。对于专题地图而言，一般使用几种不同颜色就足够了，但对于照片或具有照片真实感的图片，使用16位或者24位图像格式将更为有益。

6. 100. 大多数图像格式均采用某些压缩形式以减少文件的大小。最简单的压缩方案是长线程编码，有时一些光栅式地理信息系统也使用这种技术。如果在图像的一行中有一连串相同颜色的像素，那么系统只要存储重复的次数和记录一次该像素就可以了。例如，一连串有5个像素，其颜色编号都是4，只用两个数码代表就可以了，写成5、4，而不必记录为4、4、4、4、4。这里的颜色编码实际上只是一个索引代码，指向一个颜色表，该表保存在一个小文件头内，其中列出颜色的规定值，这类规定值总是基于某种通用模型编排的，例如RGB。

6. 101. 几种标准光栅文件格式是：

- **BMP位图格式。**微软Windows采用的不受设备制约的位图格式，几乎在任何一种显示设备上均可通过Windows显示这种格式的图像。这是一种最基本的光栅文件格式，支持长线程编码的压缩方式。尽管如此，这种格式的文件尺寸通常要比其他格式大。
- **TIFF标记图像文件格式。**这种格式是最受广泛支持的光栅图像的格式之一，支持多种不同的颜色数量和很多压缩方案。几乎所有支持图形的软件包均可导入TIFF格式的图像，尽管有时导入不同硬件平台上建立的图像时会出现一些问题。TIFF格式在地理应用方面具有特殊的重要性，在地理信息系统或桌面地图绘制软件包里显示卫星图像、航摄照片、扫描地图或其他光栅数据时，经常采用这种图像格式。由于需要有一套独立于平台的标准文件格式以便表达地球空间图像，所以制定了GeoTIFF标准。该标准提供的信息参数包括在TIFF图像的文件头内，这些参数对所有与图像有关的地理信息进行描述，例如投影法、真实世界坐标、地图范围等，同时也与标准的TIFF格式参数保持一致。GeoTIFF受到大多数地理信息系统销售商、政府部门和学术机构的支持。
- **GIF可交换图像文件格式。**这种格式设计用于跨越硬件平台进行光栅图像的交换。它所支持的压缩方案能显著地缩小文件尺寸，因此成为通过计算机网络进行交换的最佳方式。事实上，CompuServe当初开发这种格式时，是为了用于早期网上公告板的服务。GIF格式可支持256色，是网络浏览器支持进行浏览操作的两种光栅图像格式之一。大多数网页上非照片式的光栅图像均采用GIF格式。

- **JPEG联合图像专家组图像格式。**这种格式是由联合摄影专家小组作为一种图像压缩方案开发出来的，支持数量众多的颜色或灰度，可以满足摄影照片或照片式图像的需要。JPEG格式也得到网络浏览器的支持，用于在网页上显示照片。JPEG格式具有多种压缩方式。但这种压缩并非完全可逆，也就是说，经过高度压缩后的照片再也不能完全再现原始图像的所有细节。一种新的格式，JPEG2000，如今得到地理信息系统软件程序的支持。
- **PNG。**PNG(便携式网络图形)文件格式是作为GIF文件格式的免费的和开放源的接替者创建的。PNG文件格式支持真正的色彩(1 600万种颜色)，而GIF文件格式只允许256种颜色。当图像有大片单色时，PNG更好。“无损的”PNG格式最适合编辑图片，而像JPG那样“有损失的”格式最适合照片式图像的最终发布，因为文件尺寸较小。许多老式的浏览器还不支持PNG文件格式；不过，随着Internet Explorer 7的推出，所有流行的现代浏览器都充分支持PNG。Adam7-interlacing允许早期预览，即使只传输了一小部分图像数据。新的geography markup language(GML)得到了OpenGIS Consortium([www.opengis.org](http://www.opengis.org))的推荐，目前正在执行(更多信息见[www.opengeospatial.org/standards/gml](http://www.opengeospatial.org/standards/gml))。
- **DGN。**MicroStation设计文件格式(.dgn)是Bentley的Modular GIS Environment(MGE)和Geographics GIS软件包使用的。该格式不直接支持属性数据，但提供与外部数据表的链接。一个单独的导出格式将地理文件与属性文件相结合。

6. 102. 正如可提出证据证明的那样，PDF是一种独特的文件格式，因为它不适用于光栅或向量类型，但使用无疑很广泛。

- **PDF可移植文档格式。**可移植文档格式也是由Adobe公司开发的。当初的目的是为了能够在因特网上传播复合型文件，也就是说，文件中既有文本又有图形。只要安装了Adobe Acrobat打印驱动程序，就可以用它从任何一种文本处理或图形处理软件创建PDF文件。PDF阅读器可以从Adobe的网站免费下载。有的专家预言：PDF格式将会取代外部页面格式文件成为用于高级图形打印的主流标准。PDF的语言比外部页面格式简单，因此更容易对PDF文件进行光栅化处理。图形文件如果要显示在计算机屏幕上或得到高分辨率的打印效果，就必须进行光栅化处理。

### (c) 地理信息系统数据格式

6. 103. 各种格式可用于以更原始的形式向专门的地理信息系统用户和空间分析人员发布空间数据，其中包括地理坐标数据格式和表格数据格式，这些数据可作为属性与地理信息系统中的空间数据相连。

(d) 坐标数据

6.104. 由于所支持的数据格式不同,各种地理信息系统软件包之间有很大区别。每一种商业软件包都有自己的数据格式。此外,导入和导出功能使得用户可以对部分经过挑选的外部数据进行转换。有时,这些转换功能需要单独购买。在大多数情况下,许多销售商的软件可以阅读其他销售商的软件。

6.105. 尽管商业和公共地理空间技术团体都做出了努力(关于GML数据格式的最新讨论,见,例如, [www.opengis.org](http://www.opengis.org)),但至今为止仍未出现普遍接受和广泛使用的通用的或开放源的地理信息系统数据交换格式。事实上,由领先的地理信息系统销售商开发的几种交换格式成了事实上的标准,它们也得到其他软件系统的支持。以下介绍几种最重要的格式:

- **Arc/Info导出格式(.e00)**:一种跨平台的地理信息系统数据库交换格式,由环境系统研究所的Arc/Info GIS开发。导出的文件可以压缩,支持较小的文件尺寸。然而,为了保证最大的兼容性,最好使用非压缩的导出格式。导出后的文件仍可用标准压缩和存档程序(如PKZIP)进行压缩。这种.e00格式虽未公开发表,但很多地理信息系统软件包已经为其开发了导入程序。
- **ESRI的图形文件(.shp)**:ESRI的ArcView和ArcGIS桌面软件使用的一种简化格式。一个图形文件数据库由若干个文件构成,分别包括坐标数据、空间索引和属性数据。其文件格式已经发表,许多其他地理信息系统能够导出图形文件。
- **File Geodatabase格式**:一种新推荐的本地数据格式,同ArcGIS一起使用。File geodatabases以文件夹的形式储存在一个文件系统中,其设计是为了支持地理数据库的全部信息模式,包括拓扑结构、光栅目录、网络数据集、地形数据集以及地址定位器。每个File Geodatabase文件可以有1百万兆字节大。个人地理数据库使用Microsoft Access文件系统(.mdb),可以有2千兆字节大。ArcSDE空间数据库引擎也使用file geodatabase。
- **Google的KML(Keyhole Markup Language)**:一种以XML为基础的语言模式,用于表达有关现在或未来的基于网络的在线地图(2D)和Earth浏览器(3D)的地理注解以及显现。最初开发KML是供Keyhole使用的,该公司在2004年被Google收购。KML文件指定了一系列特征(地标、图像、多边形、3D模型、文本说明等),用于在Google Earth、Maps和Mobile或者执行KML编码的任何其他3D Earth浏览器(geobrowser)上显示。每一个地方都有一个经纬度。其他数据可以使察看更明确,例如倾斜、朝向和海拔,这些数据放在一起可以定义一个“相机视图”。KML文件常常作为KMZ文件发布,KMZ文件是经过ZIP格式压缩过的KML文件,带.kmz扩展名(关于GML,见上文6.101段和6.105段)。



- **MapInfo交换格式 (.mif)**：用于与MapInfo制作的文件进行交换的格式，MapInfo是一个领先的桌面制图系统。MIF文件符合ASCII格式，可以借助很多程序进行阅读。
- **AutoCAD DXF格式 (.dxf)**：起源于计算机辅助设计领域，非常适合于转换地理坐标数据，但在转换属性信息方面表现不佳。

6.106. 上述所有格式均支持边界信息和属性信息。任何一种商业地理信息系统软件程序至少能够导入其中一两种格式。在理想的情况下，一个人口普查办公室应该为公开发布的地理信息系统数据库提供若干种格式，因为用户的地理信息系统功能各异，软件平台多种多样。选择一种发布格式时，应该了解人口普查数据的广大用户群体最常使用哪些制图系统，并且要考虑数据格式的灵活性和可靠性。

6.107. 按照地理信息系统本身的内部格式(如Arc/Info或MapInfo)发布地理信息系统数据并非良策。原有格式的数据常常无法转换到其他操作系统内，可能遇到路径名不兼容的情况，而其他地理信息系统软件包通常不能导入地理信息系统数据专用格式。因此通常最后使用大多数商业地理信息系统软件采用的可靠的数据交换格式。

#### (c) 表列数据

6.108. 新的地理信息系统软件的开发不再强调表格的导入，而是强调关系型数据库，例如Oracle和Access。大多数地理信息系统软件包在属性数据方面仍然支持多种文件格式。有的软件包具备的功能可以把相关的数据库连接到外部的数据库管理系统。然而，为了发布数据，最好采用一种简单而广泛使用的数据表文件格式。应用最广泛的格式是dBase(.dbf)格式。大多数数据库管理和电子表格软件，以及人口普查表格处理软件包，例如REDATAM，即(微型计算机的用于小区域的数据检索)，以及CSPPro(即人口普查和调查处理系统)，这是一套可免费从美国人口普查局获得的人口分析工具，都采用这种格式。逗号分割数值(.csv)格式也用于表格，并且不是销售商专用的。

6.109. 虽然以dBase格式发表的表列数据能与其他地理信息系统软件广泛兼容，可它也有若干局限性。例如，字段名列于表格的第一行，限定不得超过10个字符。有关电子表格或者数据库管理的软件文献对兼容性问题进行了详细讨论。在表格布局方面，最重要的字段是通用标识符，借助标识符可以把属性数据链接到报告单位的边界上。这个字段应该位于每个属性表的第一列。将数据集按某一标准排序是一种好方法，例如按地理标识符排序。

## D. 打印

### 1. 概述

6.110. 为了便于打印少量文本或质量控制图片，人口普查办公室应当配

备一台或几台打印机。打印技术不断在发生变化，可利用的产品很多。在选择适当的打印机时，人口普查办公室需要考虑下述标准：

- 硬件、维护和打印每一页的成本。
- 吞吐量(每分钟打印多少页)。
- 以每英寸点数(dpi)和色彩或灰度数目表示的输出分辨率。
- 材质的大小。
- 所支持的材质的类型(普通纸张、特别涂层的纸张、幻灯片等等)。

## 2. 打印机类型

6.111. 以下是最流行的打印机类型(就大多数国家统计局办公室的需要而言，最具成本效益和可靠的选择是激光打印机，但还有其他选择)：

- **激光打印机**通过激光束和光学结构对光电导表面进行有选择的放电。极性相反的带电墨粉被吸附到该表面，分布在有电荷的位置上。墨粉通过转移作用在纸张面上固定下来。这种类似于静电复印的过程也可通过滚筒方式把图像加到纸面上。单色激光打印机达到的输出质量几乎与专业排版系统输出的产品不分上下。彩色激光打印机的价格直到最近才有所下降，使大多数进行图形处理的用户可以考虑使用。
- **喷墨打印机**通过一个喷嘴将带电的颜色液滴喷射在纸张表面。液态喷墨打印机采用液态墨水，经蒸发干燥、墨水经受液压由喷嘴送出。这种技术被称为脉动喷墨技术。相比之下，热喷墨技术对喷嘴处的墨水小泡进行加热，当小泡足够大之后，强行通过喷嘴射向纸张表面。固态喷墨打印机的墨水需要经过融化后才能喷射到纸张上，到达页面之后会快速固化。固态喷墨打印机在页面上形成的墨点要比液态喷墨技术所形成的墨点更为精细。喷墨打印机允许使用普通纸张，但为了达到更好的输出效果和质量，通常使用有特制涂层的专用喷墨打印纸。由于价格合理，便于操作，适用于多种幅面的输出，所以，喷墨打印机成了当前应用最为广泛的彩色输出装置。
- **热作用打印机**要求配备专用纸张和涂有墨水的专用色带，色带在热作用机头上移动。通过机头的热升华作用将墨色融入纸面。这种方式有时需要三种颜色的色带(青、品红、黄)，有时需要四种颜色的色带(青、品红、黄、黑)，因此，热作用机头需要在纸张上经过三遍至四遍。在热蜡式打印机上，通过加热把彩色蜡层附着到纸面上。在热作用的染色过程中，颜料扩散到可着色的表面上。染料扩散式打印机与热蜡式打印机相比，能获得更高的分辨率和更多的色彩。
- **静电打印机**借助静电荷将墨粉转移到非导电性的表面上，有的地方吸附墨粉，有的地方排斥墨粉。直接静电式打印机把电荷直接加到具有特别涂层的纸面上。不同颜色的墨粉需要单独的操作。所有颜色的墨

粉都走过一遍之后，墨粉将融入纸张之内。另外还有一种静电加工方法，即彩色复印。彩色复印通过曝光时生成附着性静电荷的鼓筒或带子实现。

6.112. 很多地图的样稿并不一定要用彩色打印。事实上，小幅面黑白地图更容易复印。激光打印机可打印A4或信纸大小的纸张，打印速度快并可获得很高的分辨率(600dpi或更高)。它们是打印报告和其他文件的理想手段，这些文本以文字为主，配有插图或地图。

6.113. 打印复杂地图则需使用彩色打印机，这种地图仅用单色灰度变化和符号方式难以充分表现。当前最常使用的彩色打印机为喷墨打印机，打印尺寸从A4/信纸大小一直到60x90厘米或24x36英寸的版面不等，尽管彩色激光打印机可能取代喷墨打印机。喷墨打印机可以输出600dpi的高质量地图。对于喷墨打印机来说，打印速度仍然较慢。

6.114. 对于地理信息系统项目而言，如何选定一台合适的打印机，最重要的还是费用问题。但必须记住，购买打印机的钱仅是众多费用中的一笔，而且通常是比例相对较小的一部分。有时打印机本身的价格降低了许多，但墨盒和专用打印纸的价格仍然居高不下。有时，打印机制造厂家有意把硬件设备的价格压得非常低，但他们往往希望通过销售专用消耗品获取利润。除了购买时的价格外，对打印每一标准页(例如，按页面5%的覆盖率计算)的所需成本也要进行比较。对此，可以参考计算机报刊杂志上公布的一些对比性资料。

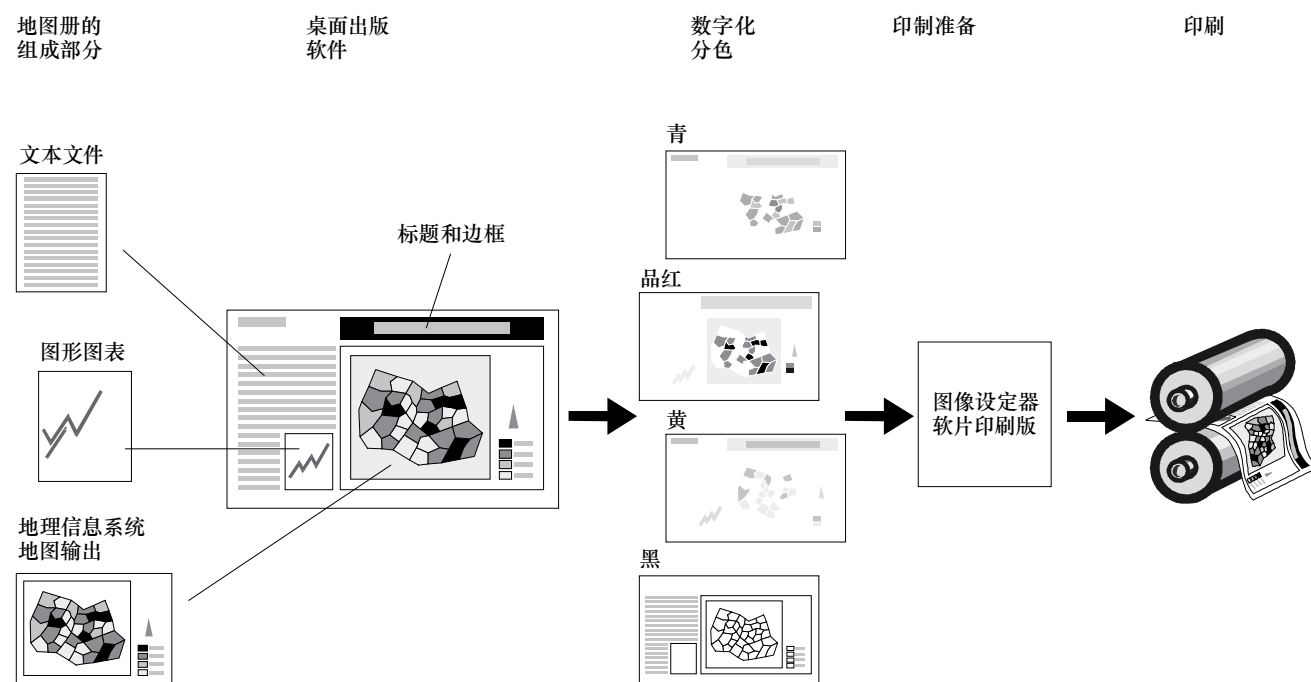
### 3. 商业化印刷

6.115. 对于大批量印制，个人打印机就显得太慢，而且成本也太高。因此，宣传册、招贴画以及人口普查地图册需要由内部的印刷车间或商业化印刷厂印制。如果印刷量相当大，采用印刷版在平板印刷机或类似印刷机的传统印刷工艺仍然要比数字化的印刷便宜和快捷。

6.116. 然而，制作印刷版之前的这段工艺，目前已经几乎全盘数字化了。数字化人口普查地图册的典型制作过程见下文(图VI.9)。在初始规划阶段，确定出文本、图片和地图等各项内容，然后由人口普查地图制作人员制成地图册内的所有地图。这些地图用外部页面文件格式保存，以便印刷。对于包含用外部软件生成的图形或照片的复杂地图，其版面设计一般要通过高端图形软件完成。人口普查办公室的其他工作人员通过标准的字处理软件为地图、表格、参考资料和其他文本内容添加文字说明。

6.117. 第二步，通过桌面出版软件对地图册中的所有组成部分进行组合，对文字标题、图表说明、图片、文本以及图形进行排版，使其组成引人注目的版面，版面的大小要与产品的页面尺寸完全相同。这项工作既可内部完成，也可拿到外面加工。在决定是否外包时要考虑的问题包括质量、成本、需要的份数和时间。

图VI.9  
数字化印刷过程



6. 118. 最后的地图册版面设计一旦完成，即可以数字化输出文件的形式加以保存。最普通的文件格式是封装式外部页面格式文件，当然，有些商业化印刷机也可使用某些软件制定的文件格式。很多高端图形软件和桌面出版系统也可进行分色处理，各分色版既可存成各自的文件，也可存入同一文件中。实际印刷机使用四色印刷版，颜色分别是青、品红、黄、黑(即所谓的CMYK颜色模型)。地图和图形上的颜色是按不同的百分比将四种颜色调和在一起产生的。随后将数字文件送至图像设定器制成软片，再从软片生成印刷版。一般来讲，利用数字文件制作软片效果是最好的。激光打印机上打印出来的照相制版拷贝可通过翻拍技术复制，尽管便宜一些，但无法达到数字文件所具有的高分辨率。除非现有的生产线已经过调试，一般都要先让印刷机出一份彩色清样，对色彩进行评估后方可最后生产。很多印刷硬件和软件的销售商也都通过各自的网站提供广泛信息以及其他资源。

### E. 用于发布目的的数字地理数据

6. 119. 随着软件越来越成熟，因特网在全球的应用更广泛，发布数字地理数据也成为可行。对数字化数据库的需求将与日俱增，该数据库中包含人口普查部门的数字化地理主数据库中的内容。在很多领域进行政策制定和科学分析时，人口普查数据是一种重要资源。医疗保健服务的提供、教育资源的分配、公用设施和基础设施的设计以及选举计划等就是一些应用实例，政府有关部门在开展这些工作时，需要参考小区域的空间基准人口统计资料。商家也会根据这些数据进行市场策划和选择恰当的销售地点。

6. 120. 数据处理和地理信息系统方面的进步意味着在点查区或类似的分解层面建立一个数字化地理数据库对许多国家统计办公室来说越来越容易。为了满足用户需要，国家统计办公室应遵循一个数字化数据发布策略。国家统计办公室应考虑提供如此详细的数据的成本和收益。

6. 121. 好处包括极其详细和精确；有可能将人口普查数据用于多种用途，尤其是在覆盖到其他地理数据如地形数据上时；几千个单位的管理和储存相对容易。

6. 122. 成本包括数据处理和管理的费用，数据披露可能产生的问题以及质量控制。在数据收集单位的层面发布人口普查结果可能会遇到在人口普查过程中不曾预料到的错误。元数据必须始终伴随地理数据的发布，元数据的制作成本也应考虑在内。

6. 123. 除发布点查区数据库外，另一个选择是制作一个详细程度类似的衍生产品。

## 1. 针对潜在用户的数字化数据发布策略

6. 124. 正如上文所指出的，国家统计办公室在规划其产品和服务时应考虑更广泛的数据发布的成本和收益。小区域人口普查数据有广大的潜在用户，这意味着国家统计办公室需要推行一个多层次的数字化数据的发布战略。一般来说，我们可以区分以下类型的用户：

- **地理信息系统的高级用户。**他们可以轻易地操作大型数据集并且可以使用ftp(文件传输协议)来获取这些数据。他们需要广泛的元数据。高级用户通常将小区域人口普查数据与他们自己的关于保健设施、学校区或销售区等的地理信息系统相结合。这些用户有时被称为数据提取者或“强势用户”。他们想要获取完全数字化地理格式的空间信息和属性信息。人口普查办公室需要提供关于用于地理数据库的地理参数以及关于个别的人口普查变量的完整文献(见关于元数据的开发的第三章F节)。空间信息将以开放的地理格式发布，这种格式可以很容易地转变为任何商业地理信息系统格式。
- **政府、工商界或私营部门中的计算机用户。**他们需要广泛浏览人口普查数据库中的专题信息，需要能生成专题地图，需要对绘图参数进行简单处理。此外，他们可能还需要利用一些简单的分析功能，例如，把人口普查的单位纳入到用户定制的区域内。最好向这类用户提供一个全面的、预先打包、为商业的或免费的桌面绘图软件设计的应用软件。文件要求较少，因为用户不太可能改变数据库的地理参数或者执行更高级的地理信息系统操作。
- **初级用户。**他们大多只是想在计算机上看一看现成的地图，或许进行一些简单的查询。对于这类用户，最好的数据发布策略是制作一个独立的数字化人口普查地图集。该地图集可以包括一系列静态地图，例

如，幻灯片格式的。或者它可以是非常简单的制图界面，带有事先设计好的地图，可以对这些地图做基本的查询。不论是静态地图还是一个简单的地图界面都可以通过因特网访问。

(a) 数据内容的定义

6. 125. 为了发布地理信息系统数据库，第一步的工作就是对数据库的内容进行定义。国家统计局办公室应规划单个数据集的发布，而不是用于内部管理的可操作的数据集。这样做的一个主要原因是，把两类数据分开将使得国家统计局办公室可以在持续的基础上继续维护和更新其内部数据。为了发布数据集，需要解决许多问题，比如下面列出的这些。

(i) 发布何种层次的数据？

6. 126. 为了最大限度地发挥人口普查数据收集工作的总体效益，人口普查机构的目标应该是，在不妨碍数据保密性或统计有效性的情况下，发布最基本一级的地理基准人口普查数据。即使在点查区一级，仍可能包括只有几家住户的特殊报告区，在这种情况下，则不能公开人口普查数据。如确有需要，选择的报告区数据必须删除或者予以重新编码。

(ii) 是一个大型地理信息系统数据库，  
还是一系列人口普查数据库？

6. 127. 一个高分辨率的人口普查地理信息系统数据库包括数以千计的报告单位。如此大的数据量将超出一般普通数据用户的计算能力。此外，国家统计局办公室应考虑数据文件的实际大小以及下载文件对服务器的影响。例如，利用一个56千字节服务器下载一个3千兆字节的数据文件需要3天的时间。

6. 128. 在这种情况下，人口普查机构应考虑制作一系列人口普查数据库，而不要发布一个大型数据库。就中等分辨率的数据库(例如区域性数据库)而言，一个国家级的摘要式数据库即可满足提供国内社会经济状况信息的需求。对于主要地区或区域，可以建立独立的、显示分区和点查区一级各项指标的数据库。对于主要城市地区来说，建立独立的数据库也十分有用。

6. 129. 最后，带有相关人口普查数据的国内居民点的数据库会满足一部分用户的需要，他们不需要报告单位的地理信息系统数据库的空间分辨率。这种数据库至少要包括按城市地区分类的居民区，以及每个城镇的加总人口普查指标。理想的情况是，把村庄一级的数据库建立起来，这样会给卫生健康、教育或农业部门的规划者带来很大益处。村庄一级的数据库可以地名字典为基础，如果点查前实地工作期间已经收集了有关地名和位置的信息的话。

6. 130. 为地区一级提供数据库将会增加数据的用途。很多用户仅仅需要相对较小的区域的人口普查信息，通常是他们自己的区域的普查信息。对于那些具有中等地理信息系计算能力的用户来说，处理国家人口普查数据库的子

集相对比较容易。对于那些数据访问费用高于复制成本的国家来说，应以大量非商业用户能够支付得起的价格，向他们提供一些较小的数据集。

6. 131. 如果发布独立的数据库，要注意各个部分必须相互兼容。也就是说，各数据库的共享边界必须严格匹配。数据库的独立片断必须处于同一个地理基准系统中，具有同样的数据库属性定义。如果人口普查办公室的主数据库非常详细，向某些用户提供一个更通用的数字化普查地图版本将大有裨益。有些国家提供的数字化人口普查地图使用了不同的标称比例尺或坐标准确度。对于有高准确度和详细要求的用户可提高收费。

6. 132. 很多商业化地理信息系统数据的制作没有按照特殊投影，而是以经纬度(地理坐标)发布他们的数据。地理坐标是最为普通的基准体系，如果用户需要与其他数据层共同使用人口普查边界数据，可以很容易地将地理坐标转换导入其他投影系统。相比之下，地理信息系统软件一般不支持某些国家专用的投影和坐标系统。因此，用户在这种情况下进行地理分析应用时，很难使用普查数据库。

(iii) 如何确定边界与数据库的集成密集度?

6. 133. 人口普查地理信息系统数据库的特点是具有大量的属性字段。人口普查问卷上的信息可能要存储到数百个变量字段内，一般不可能将其放入同一张数据表中。一种比较好的办法是，在地理属性表内选择一些最重要的指标，把其余信息量分列成一系列不同表格。外部表格可以按不同的题目排列，如人口统计、住户数据等等。需要时，用户可以利用通用地理识别符将表格链接到地理信息系统上。必需再次强调编码的一致性，以确保以惟一的方式定义各种地理实体。

(iv) 需要提供多少元数据?

6. 134. 虽然对任何数据发布来说，元数据的生成和文档作为伴随物是绝对必需的，但所需信息的具体数量随用户而变化。一般需要大量元数据的内部用户与一般不需要那么多元数据的外部用户之间存在重大区别。见本章后面的部分以及前面的部分(关于元数据的更多信息，见上文第二章和第三章F部分以及下文第6.137-6.139段)。

(b) 文件命名约定

6. 135. 虽然Windows、Macintosh、UNIX和LINUX操作系统都支持长文件名，然而，对于所有数据和文档文件的发布最好还是遵照DOS 8.3(即8字符)文件命名约定并为保持一致性制定内部标准。某些用户可能仍在使用DOS或Windows 3.1或老式地理信息系统软件包。短文件名可以减少不兼容问题，例如，与老式网络软件的不兼容问题。在文件命名方面使用一致约定并在文档中说明，便于用户迅速找到所需的数据。

(c) 压缩

6. 136. 地理信息系统的文件一般非常大，而且还包括表格数据，发布一整套文件将是一项巨大的工作。对文件进行压缩可为数据发布提供很大方便，特别是通过因特网或数据磁盘或可读光盘发布时尤其如此。在Windows环境下使用最广泛的压缩软件程序是PKZIP和Winzip。几乎所有计算机上都可以找到这两个程序。在UNIX操作系统下也有从压缩文件夹中解压缩文件的程序。不过，应该注意的是，一些格式，包括File Geodatabase，不能压缩。自解压文件对于没有经验的用户更为方便，因为无需使用解压的程序。然而，它们要在特定的操作系统上运行，只有在知道目标平台的情况下，才能使用这种自解压文件。

(d) 文档，包括数据字典

6. 137. 与数据集一起发布的文档不一定要像人口普查机构内的信息那样完备，机构内部的信息是为所有数据库进行编制的(见上文第二至第五章)。任何用户都可以阅读简单ASCII(“read-me”)文本文件。数据用户通常不必详细了解数据的系属来源或处理过程，重要的是要让外部用户容易理解和使用。因此，文档应该针对数据库中与学生密切相关的方面做出清晰、简洁和详细的说明。只要人口普查办公室拥有一个全面的元数据库，为用户提供的数据文档就可以很快编制出来。数据文档应包括以下信息：

- 数据集的名称和基准信息，包括所有数据来源。
- 数据集的解说内容。
- 行政部门与报告单位隶属关系及其与其他地理特征(如居民点)的关系的说明，包括每一类型的报告单位采用的统计定义的清楚说明。一份完整的、包括所有报告单位及其地理编码的清单十分有用。
- 软件和硬件要求。
- 通用数据格式，解压与安装说明。
- 地理基准信息(所有地理数据集必须在同一个基准系统中):
  - 地图绘制投影方法及所需全部参数，如标准的平行线或子午线、伪东向偏移量和伪西向偏移量等。
  - 坐标单位(例如十进制、米、英尺)。
  - 源地图比例尺(硬拷贝地图的比例尺，根据该比例尺对边界进行数字化处理)。
  - 地理准确度信息(例如，对源地图是否有数量准确度的信息可供报告)。如果无法对数据质量做出量化评估，也可用一般性词汇对准确度进行说明。
  - 地理信息系统数据集的印刷版地图是对文档的有益补充。例如，它能检验用户导入地图的操作是否正确。
- 与互不相连的报告单位沟通的约定。



- 有关产品信息，例如，更详细的人口普查地理信息系统数据库，或可用于边界的额外数据文件。
- 有关人口普查的出版物的参考目录。
- 用于用户支持的联系信息。
- 免责声明、版权信息，等等。

6. 138. 另外，每一个地理信息系统数据集应该配备一部数据字典，其中载有每一个地理信息系统数据层或数据表的信息。“数据字典”是一个比较老的术语，指的是数据集里的特定格式化和字段名。应该列出的信息包括：

- 文件名和文件格式。
- 特征类型(点、线、多边形)。
- 坐标数据文件与相关外部属性数据表之间的关系。
- 要在属性表内和附加的外部表内的每一个字段中列出：
  - 字段名称。
  - 字段内容描述(例如，人口总数，2005年)以及采用的精确统计定义。对于推导出的人口统计指标，应指出所使用的公式，例如，分子和分母变量使用的字段名。
  - 字段定义，包括变量类型(例如，实数、整数、字符字段)，可接受的数值范围以及处理脱漏数值的约定。对于分类数据，编码方案要详细解释。例如，在一个居民点数据库内，一个叫做TYPE的数字字段，可以用“1”代表国家的首都，用“2”代表省会，用“3”代表区的行政中心，等等。
  - 可用的数据质量信息，用户可以通过这一信息判断数据是否适合于某项特定任务。

6. 139. 数据文档和数据字典也可以合并为一本全面的用户手册。用户手册可以包括有关数据库内容、数据的谱系和数据质量方面的详尽说明。此外，还可以逐步说明应用实例或用数据库创建的人口普查地图的范本。数据字典的样本见附件四。

(e) 可交付的数据产品的质量控制和保证

6. 140. 如上所述，质量控制是最终成品发布前的一个重要步骤。在数据库最终版本制作完成，并且做成了将要发行的相应形式(例如压缩)之后，应在所有目标平台上对数据库进行测试(例如Windows环境、UNIX、Macintosh、LINUX)。

6. 141. 对许多用户来说，只读光盘仍是用于发布大型数据集最合适的介质。一张只读光盘可保存700MB的内容，而大多数计算机都有光盘驱动器。光盘刻录机也不是太昂贵，因此可在单位内部制作数字化母盘。这有助于定制化数据集的发布，因为用户通常只需要少数几份拷贝。对于大范围发行的大型数

据集，只读光盘有不少优越之处，如单位成本低，经久耐用，可在多种硬件平台上读取等。

6. 142. DVD(数字化视频光盘)技术在某些领域取代了只读光盘，并且DVD刻录机在台式电脑上越来越常见。一张通用的、单面的、标准的DVD可以保存4.7吉字节的信息。

6. 143. 更长远地看，大多数数据发布将会通过因特网完成。当前，限于带宽，在许多国家非常大型的文件在发布上仍然受到阻碍。所谓带宽系指在一定时间内所能传输的数据总量。在许多国家，下载时间太长，常常令人难以忍受，其中部分原因是因特网基础设施不足。主要的瓶颈是家庭或办公室用于连接因特网的调制解调器。对于那些配有高速因特网访问能力的科研部门、政府机关或商业用户来说，大型文件的传送已不成问题。

6. 144. 对于人口普查部门来说，通过因特网发布数据将大幅削减再制作成本。余下的只是软件界面开发、网站维护以及逐渐增多的网络服务器资源占用费。届时，可以以非常低的收费，甚至免费向用户提供人口普查地理数据库。但某些组织可能会收取在线数据使用费。收费的一个理由是为了补贴出书费用，为没有条件上网的用户提供出版物。另一个原因是，人口普查机构需要收回数据收集和人口普查数据编制的部分费用。

## 2. 法律和商业问题

6. 145. 下文论述国家统计局办公室应该考虑的与地理数据的生成和拥有有关的问题。这些问题包括数据版权、地理数据商业化与义务的平衡。

### (a) 数据版权

6. 146. 版权是一种专有的和受法律保护的出版、复制或销售一件作品的权利，此处的作品就是数字地理数据库。由于数字数据非常容易复制，因此，与印刷式地图相比，地理数据库的版权问题就成为了一个迫切需要解决的问题。人口普查办公室需要制定一个用于表格和图形形式的普查信息的数据访问政策。

6. 147. 版权涉及两个领域：道义上的权利和物质上的权利。道义权利保护作品的完整性，禁止对原作进行任意改动。物质权利指在发行作品供复制、使用或转换时获取金钱利益的权利。版权拥有者的权利将以许可协议的方式予以确定。

6. 148. 版权问题涉及数字化数据产品的定价政策。在制定数字化空间数据的价格时，人口普查机构有几个选择。人口普查机构可以决定：

- 自行承担全部人口普查数据的收集和发布费用。
- 收取数据发布使用费(载体介质和发运费)。
- 全部或部分收回数据收集和编制费用。

- 除了收回地理信息系统的投资和实际数据开发成本外，进行创收。

(b) 地理数据商业化的权衡

6. 149. 不同国家的版权法有很多差别。最极端的情况是，有些政府对公共部门制作的信息产品不拥有版权。他们的根据是，纳税人已经为数据收集工作付出了他们应缴的税款，因此不应再为使用这些数据而交费。这样做结果是，公共部门免费提供他们所制作的地理数据或只收取复制费用。此外，任何商业机构可以使用政府的信息产品，对产品进行再包装，以营利为目的出售。

6. 150. 例如，在美国，人们可以免费访问公共数据，因此出现了一个庞大的服务行业，该行业以各种格式生产空间基准人口普查数据，用于向个人用户、商业用户和公共部门用户销售。尽管公司对这些数据收费，但这种开放式的人口普查数据使用方式仍使大批公司进入市场。竞争降低了经过再包装的人口普查数据的价格，同时增加了专用产品的范围。那些愿意自己动手转换数据的用户仍可继续免费使用各种数据。

6. 151. 这种发展的好处是，人口普查数据在地理方面得到了非常广泛的应用。用户数量的日益增加，促进了易用型桌面地图绘制软件的开发和增值服务的开展。这一发展取得了全面的经济效益，因为增加了税收，信息获取机会的增加导致了生产率的提高以及公共和私营部门决策的改善。这些收益足以证明数据免费发放政策的正确性，这相当于政府为私营公司提供了补贴。

6. 152. 在其他一些国家里，政府紧缩开支，对公共部门创收以支持部门运行造成了巨大压力。在地理基准普查信息方面造成的结果是，信息价格有时非常高，因而限制了信息的使用。高昂的价格也许反映出这些数据所具有的商业价值，例如对金融机构和各种企业所具有的价值。但这样一来，小公司和非商业用户不得不退出普查信息市场，限制了人口普查数据的全面应用，结果减少了人口普查地理数据的效益。正如Prevost和Gilruth(1997年)指出的那样，企图回收成本的努力使非商业用户无力承担人口普查地理产品的费用，从而导致非法复制数据集、花费时间根据原始材料重新进行数据开发或使用一些廉价的劣质数据等现象发生。

6. 153. 限制性的许可协议也阻碍了人口普查产品的发行和服务的提供。这种做法降低了人口普查数据收集的公共福利作用。由于缺乏这种附加作用，人口普查部门的收入也许有所增加，但减少了国民经济的整体收入。事实上，在有些国家，由于认识到高收费不但没有起到强化版权的作用，而且由于重要信息使用不足造成了社会效益的损失，因此恢复了免费或降低收费的发行政策。

6. 154. 当人口普查办公室与私人数据制作商合作，或公共部门或私营部门数据制作商的数据被用于人口普查地图的制作时，常常在数据访问和二次利用方面受到很严格的限制。例如，人口普查机关与一家私营地图绘制公司签订

了由私营公司承担人口普查数字地图制作的部分费用的协议，但私人公司只有获得了地理数据的专有销售权后，才能收回投资(当然，这种情况对于经销商来说不存在任何问题，他们只是从私营公司那里购买服务，而全部产品仍然是人口普查机构的财产)。

6. 155. 如果制作人口普查地图的数据来自其他部门，如国家制图机构或地方政府时，人口普查地图上应详细标明定价、版权、数据来源以及资信信息。应尽量避免版权方面的冲突，当人口普查机关将来在人口普查制图活动中需要与这些部门合作时尤其如此。

6. 156. 在大多数国家中，有一种权衡折中的处理办法，既不是最大程度地全部开放人口普查数据的访问，也不是对人口普查数据采取高收费政策。他们从两种极端做法中找到了妥协方案。例如，在那些愿意将彼此的数据纳入到对方产品中的各个政府部门之间进行特殊安排。人口普查机构可以与国家测绘组织达成协议，将道路、河流等数字化底图发送给人口普查地理信息系统数据的用户。另外，对于科研机构或其他非营利部门的用户给予折扣。还有一种办法是，只免费提供某些一般性产品，而对需要更多处理和有更多附加值的产品则采取收费政策。

### (c) 责任问题

6. 157. 法院曾在若干案例中做出判决，如果由于错误地理信息导致事故或其他伤害，地理数据的制作人要负法律责任。到目前为止，很多案例所涉及的故事多是由于地形图上缺少必要信息或出现错误造成的。根据记录的案例，导航图上的错误曾导致飞机坠毁和海上事故。地图的设计和 Information 内容取决于使用意图，然而，有时地图的使用超出了地图数据制作者的预期范围。例如，一个人口普查部门为报告单位提供数据时，可能附上街道网络的数据库。对于人口普查数据使用来说，道路信息并不很重要，在质量控制上，不如用于紧急救援的路线图严格。如果把这些有待完美的数据用于非预期的使用目的，自然会出现事故。

6. 158. 另外一个与责任问题有关的例子与人口普查数据的发布密切相关，这就是，信息的保密性是否会受到侵害。一般情况下，人口普查机构发表的都是加总数据，不会泄露个人、住户或小团体的信息。如果人口普查部门就若干小区域的地理条件重新加总微观数据，通过巧妙的地理信息系统操作就可能把低于最低披露级别的群体的信息分离出来，这类小区域可能是点查区、邮政区、医疗保健区或教育区(见上文关于区差问题的第6.32-6.36段)。在某些国家，这样做可能会成为当事人提起法律诉讼的理由。

6. 159. 有趣的是，Johnson和Onsrud(1995年)认为，销售地理信息系统数据和限制二次使用数据可增强数据提供商的责任感。收费意味着数据提供商进行了担保，保证材料没有错误并且适用于预期目的。相比之下，把数据放到公共领域，会使人口普查机关免受索赔的威胁。

6. 160. 在发布空间基准数据之前，人口普查部门应向法律专家咨询，拟定一份免责声明，与数据产品一起发行。免责声明可以包括下列要点：

- 声明信息在收集之时被认为是准确的，并且是从可靠来源收集得来的，但不对其准确性进行任何担保。
- 告诫人们，信息可能发生变化，恕不另行通知。
- 如果地理数据库的任何组成部分由外单位创建，应该明确说明。
- 指出一旦使用本数据库就意味着接受本免责声明和有关协议。

### 3. 通过因特网制图

6. 161. 很多国家的统计机构已经把因特网视为一种发布信息和数据的手段。网页的内容十分广泛，从人口普查结果的简单列表，到用户可以对专用表格进行复杂的交叉查询的界面。

6. 162. 出于多种原因，鼓励利用因特网来发布数据。不过，选择利用因特网发布数据就意味着国家统计局办公室要进行投资，尤其是为了因特网发布用于计算机服务器、存储、更新和备份上的加载。鼓励国家统计局办公室考虑作为全面的人口普查规划的一部分投资于因特网部署的总的存在期成本。

6. 163. 因特网特别适合于展示和发布地理信息。最简单的方法就是把统计办公室的静态地图一幅一幅地展示出来。例如，展示人口普查变量的系列地图可以通过桌面地图绘制软件包制作。大多数软件都允许使用标准图像格式，例如GIF或JPEG格式等，用户可以利用其中一种格式把自己的地图文件储存起来，然后就可以像别的图形和照片一样集成到网页上。这样的网站可以为数据用户提供有用的信息，但不允许他们处理数据，或为某一特定地区定制地图。以下各节概述允许用户与人口普查地理数据库之间有效程度交互式操作的方法。

6. 164. 大多数地理信息系统和桌面制图软件公司都通过标准的数据交换协议，自行开发独立于操作平台的因特网制图工具。统计部门通过此类工具可以把地理信息置于服务器上，用户可以利用标准因特网浏览器以交互式的方式对这些数据进行制图和查询。因特网用户不必购买专用的地理信息系统软件，而是在因特网上直接访问地理信息系统数据。凡是通过地理信息系统存储和处理的数据，均可以通过这种方式发布，其中包括向量地图、光栅图像和数据表。

6. 165. 把因特网地图处理软件作为一种内部工具也十分有用，统计部门的工作人员可以通过局域网访问空间数据。他们不必购买中心服务器上商业地理信息系统软件的网站许可，而是通过浏览器软件访问的所有地理信息。

6. 166. 侧重服务的结构是对包括侧重服务器模式、侧重客户模式和混合模式在内的方法的统称(下文对每种方法将做更详细的介绍)：

- 侧重服务器的策略。用户向存储数据库的服务器发出地图处理请求。服务器上的制图软件对请求进行处理，将制好的地图(比如用GIF格式)发给用户。
- 侧重客户的策略。与上述过程相反，大多数的处理工作在用户(客户)的计算机上完成。
- 混合模式，即把侧重服务器的策略与侧重客户的策略结合在一起。

(a) 侧重服务器的模式

6. 167. 这种方式有时被称为“薄客户/厚服务器”结构，它把大多数的数据处理任务交给了数据发布机构内的服务器。这一点类似于传统的主机终端结构方式，即一台强大的中央计算机为一定数量的用户进行数据管理、储存和处理，而用户通过一台处理功能薄弱的终端与中心计算机相连。

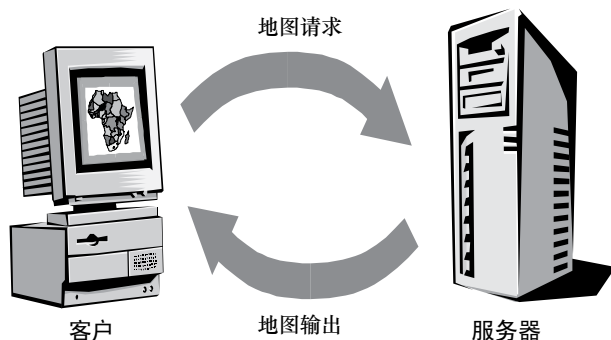
6. 168. 图VI.10简单地说明了侧重服务器策略的原理。用户连接到网站，并发出一项地图请求。用户要么通过地区或区域名称，要么通过坐标指定一块矩形，确定出所需的地域范围，然后确定地图输出的参数，包括绘制地图时的变量、分类与色彩方案，以及能够提供地貌环境(如道路、河流或行政区边界)的其他数据层。

6. 169. 用户的请求通过因特网送至服务器，经路由到达地理信息系统软件包。地理信息系统软件可直接放在网络服务器上或放在与服务器连接的另一台计算机上，地理信息系统软件既可是因特网地图绘制商品软件，也可根据需要通过商品模块定制而成。该软件从数据库中提取所需数据，制成地图，然后以网页形式送回给用户。由于网络浏览器无法处理向量数据格式，因此一般以GIF或者JPEG格式将地图作为标准图形发出。用户如果需要对这幅地图的设计进行修改，必须重新向服务器申请。

6. 170. 侧重服务器的模式的优点是，用户不需要配备功能强大的计算机来访问可能非常大的地理数据库。如果服务器有足够的容量，即使相当复杂的

图VI.10

因特网制图：侧重服务器的模式



地理空间处理过程，例如地址匹配或网络路由，也可以很快完成。对于用户来说，只需要有一个基本的因特网浏览器和因特网连接就可以了。数据完整性得到保留，因为用户无法操作数据库本身。用户还总能得到最新信息。数据提供者可以更多地控制用户看到的内容和查看方式。可以预先设置地图绘制版面，确保非专业用户也能得到可接受的地图输出。这种模式的缺点是网络访问高峰期，服务器繁忙。

(b) 侧重客户的模式

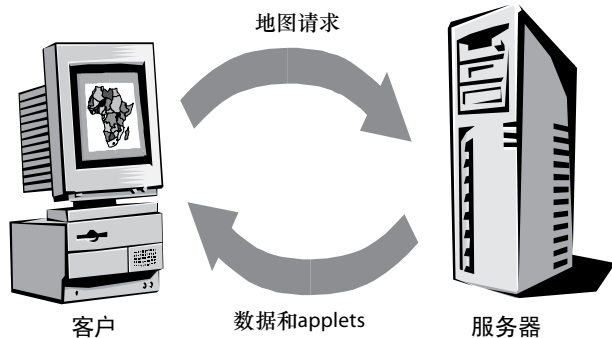
6.171. 侧重客户的模式也称为厚客户结构，即把需要处理的大量工作转移到用户的计算机上。服务器的主要任务是保存数据库和向用户发送数据库内的所需数据，有时可能还要发送一些制图处理模块。侧重客户的模式还可分为以下两种。

6.172. 第一种情况，用户计算机不具备制图处理能力。用户发出请求后，服务器在送回地理数据时，还会发送一个小程序或applet程序，用于地图绘制或地理分析(图VI.11)。applet是一段与平台无关的程序，用Java编程语言写成，可以在标准的网络浏览器上执行。收到这些之后，用户即可脱离服务器，独立处理这些数据。在浏览地图的各个图层或改变地图设计时，不需要向服务器重新请求。

6.173. 在第二种侧重客户的模式中，用户的计算机上有制图软件、applet程序或浏览器插件。插件系指一种可以扩展因特网浏览器的工作能力的程序，例如使浏览器能够显示特定格式的文件。这种方式的优点在于每次用户访问地图服务器时不必下载地图处理软件。

6.174. 下载完数据和程序之后，用户无需再与地图服务器进行更多的通信。地图处理或分析可脱机进行。可以利用用户计算机上的资源，加快处理过程。在分析和显示空间数据方面，可以给予用户更多的灵活性和自由度。不过，数据和程序文件可能非常大，需要有快速的因特网连接能力，因此，计算机功能不太强大的用户可能无法执行更复杂的制图和分析任务。侧重客户的模式允许用

图VI.11  
因特网制图：侧重客户的模式



户存储一些原始的地理数据，向服务器发出请求后就可下载到计算机上。如果统计机构服务器上的地理数据全部或部分拥有版权，则会出现问题。

(c) 混合模式

6. 175. 侧重服务器模式的优越性在于，可将相对简单的地图提供给大批非专业用户，适于普查办公室向公众展示人口普查地图。而侧重客户的模式特别适合于使用局域网的少数用户，他们具有相对丰富的地理信息系统和地图绘制方面的知识，有使用复杂数据库的能力。因此，侧重客户的模式适合人口普查办公室的人员对地理数据进行访问。

6. 176. 混合模式结合了侧重客户的模式和侧重服务器的模式的优点。它可在本地基础上为用户提供地图查询和操作方面的灵活性，但把大多数处理和分析工作交给服务器。这种模式要求客户和服务器之间在处理能力方面具有一定的通信能力。

(d) 人口普查数据发布面临的机遇，包括MapServer

6. 177. 目前，有许多可供选择的因特网制图软件。数据提供者可以购买支持标准数据集的现成软件。由于人口普查数据制图软件属于一种相当标准的应用，国家统计局不难找到适合的解决办法。在比较复杂的应用方面，需要一套软件模块工具箱，供数据提供者自行设计地图服务器的界面。

6. 178. MapServer概念是作为一种有利于打造空间化因特网应用程序的开放源开发环境设计的。Mapserver是今天可以利用的使用最多的基于网络的开放源服务器技术(见[www.opengeospatial.org](http://www.opengeospatial.org))，是将地理信息系统数据转化为网络客户观看的地图图像所需的核心部分之一。事实上，MapServer支持Open GIS Consortium(OGC)标准和特征的高级地图输出，包括专题地图的绘制和地图要素的自动化；取决于比例的特征提取和特征标记；可定制的、由模板驱动的输出；支持流行脚本和开发环境；支持地图投影；以及大量光栅和向量格式(更多信息见<http://mapserver.gis.umn.edu>)。

6. 179. 随着网络能力的提高，更大的数据集和程序模块可发送给用户，更多的用户可同时得到服务。侧重客户和侧重服务器方面的问题可随着因特网的连接速度更快而得到解决。客户计算机可在没有延迟的情况下与服务器频繁通信，使得用户的请求几乎立即就能得到满足。

6. 180. 对于人口普查数据来说，最佳的因特网数据访问和发布策略取决于用户自身的能力和专业知识。一个灵活的系统将为任何观众层次的用户提供服务，包括“强势用户”和主动和被动用户。这些类别与上文所述的用户分类大致相同。

- **强势用户**想获取整个数据库，安装在自己的使用商业地理信息系统软件的计算机上。对这类用户，要用常规的数据发布方式，例如购买



CD-ROM、DVD或从因特网上下载“粗加工”的人口普查地理信息系统数据集。

- **主动的用户**具有一定的地理信息系统专业知识，但不具备地理信息系统的本地处理能力。这些用户想要下载部分数据库，以及一些能够执行任务时需要的程序模块(applets)。
- **被动的用户**只想获得预先设计好的地图。这类用户的请求由服务器进行处理，然后通过因特网以一种合适的文件格式将最终信息送给用户。可供选择的文件格式有很多种，例如适合于地图的光栅式图像文件或外部页面格式文件，以及适合于数据的电子表格或数据库文件。

6. 181. 在因特网上，一个灵活的人口普查数据发布系统应具有以下特点：

- 用户确定自己感兴趣的地理区域，以下载数据的方式完成，或直接申请一幅地图。所需的地域范围可通过以下地址确定：
  - 地理区域名称，例如城市名、地区名或省份名。
  - 按地理坐标确定出的矩形。
  - 用户利用浏览和缩放功能通过交互操作确定的地区。例如，可以由一幅全国地图开始，将所需区域放大，然后在屏幕上画矩形或多边形，选出特定的地理区域。随着用户的放大操作，可以在地图上看到更多的细节。起初，地图只显示国家和省份的边界。用户对一个省份进行放大后，出现下一级的边界和城市位置。选中某一城市，随后出现主要街道和市内点查区的边界。层层显示取决于地图的比例尺，该比例尺对应于当前地图在用户屏幕上所占的范围。
  - 通过地理查询方式定义区域。例如，一个需要了解潜在客户所在地区的人口特点的商业用户可能需要知道这样的信息：在以购物中心所在地为圆心，半径为五公里的区域内有多少人居住。政府规划部门可能需要了解一条即将建造的公路两侧五公里范围内有多少人口。

6. 182. 用户规定所需变量和输出类型。选择包括用户能够规定地图的基本版式，例如种类数量、分类方式和灰度或颜色。或者，输出可以是一个简单的数据表显示为该地区选定的变量。用户还可以决定是否要进行数据库和地理查询，是否需要分析模块，或者确定结果是一幅地图还是数据库。

6. 183. 数据库服务器对用户的请求进行解释，并创建一个相应的数据库子集。对于根据地理名称指定的地区，这一过程只是进行逻辑选择，例如选定某一特定地区内的所有点查区。如果请求的区域名称与标准人口普查地理层次不匹配，则需进一步处理。在有些国家，居住单位的地理信息系统数据库已经建立或正在建立，所有住宅均以地理坐标联系在一起。服务器上的地理信息系统可以通过选择用户指定区域内的所有住户，编制出用户表格。如果无法完成这样的操作，服务器上的地理空间数据库则需执行一次地域插值运算，

有关技术见上文所述。

6. 184. 查询结果被发回给用户，发回的结果可以是基本数据，用户可通过地理信息系统的applet程序对其做进一步处理；或者发回一幅地图或数据库报告，供用户直接使用。当然，除了数据库或地图之外，还应提供数据文档以及其他相关信息。

6. 185. 由于各国有不同的数据发布政策，这些服务可能免费，也可能收费。如果只需得到一些已经编制完成的基本信息，则可免费提供，如果需要比较复杂的信息，则可能收费。

6. 186. 如果定制的表格以微观数据为基础，那么数据私密性是一个重要的考虑事项。与因特网的商业化应用一样，因特网的安全问题在网络人口普查数据管理方面具有十分重大的意义。对于允许访问人口普查微观数据的内部网络，必须用防火墙将其与允许外部用户访问加总人口普查数据的因特网领域隔离。

6. 187. 显而易见，未来的数据发布界面会有长足的发展。如果因特网进入每个家庭、企业和政府部门，必须有快速的因特网连接手段才能满足大量用户的需求。很多国家目前尚未达到这种水平，但由于技术的快速普及，用不了多久，很多国家均会出现通过因特网满足大多数数据请求的局面。有些人口普查机构奉行积极的数据发布策略，这些策略包括了文本探讨的一些重要因素。

#### 方框VI.1

#### 为发布人口普查数据进行网络制图的案例研究：加拿大

自1990年代末期以来，因特网一直是加拿大统计局重要的发布媒介。以网络为基础发布数据拥有巨大的潜力，可以在获取统计数据方面让用户拥有很大的灵活度。发挥此潜力的一个关键问题是提供工具，让用户可以很容易地发掘可用的数据。加拿大统计局开发了一个以网络为基础的绘图工具，可以让用户通过与地图的互动从地理的角度探索人口普查数据。有了GeoSearch，用户就能寻找、确定、显示和获取加拿大各地基本的地理和人口数据。

#### GeoSearch2006



为了找到加拿大的某个地方，用户可以与位于屏幕左手边的一套通用的制图工具进行互动，进行缩放，围绕地图运动或者确定他们感兴趣的地方。地图窗口下的一系列制表符是通向各种功能的大门，例如选择一个标准的地理区域类别，利用各种信息(地名、地址或邮政编码)搜寻一个感兴趣的区域，利用人口普查数据绘制专题地图，确定地理结构关系或获取补充数据。一旦确定了一个区域，用户可以查看人口和住房统计，从超过35个人口普查变量中进行选择以制作动态的专题地图，连接超过10 000张PDF格式的基准地图或者获取2006年社区概况中的大量人口和社会经济数据。

除了GeoSearch，加拿大统计局还提供一整套人口普查地理数据产品，可以从加拿大统计局的网站获取这些产品。这些产品分为三大系列。地图产品系列(PDF格式)包括说明标准地理区域范围的基准地图和显示与多个专题有关的统计数据的主题地图。空间信息产品系列包括所有标准地理区域的边界文件、道路网文件和水文地理数据文件。这些空间产品使用户可以建立自己的地理空间数据库，用于分析目的或将人口普查地理区域纳入它们现有的地理空间数据库。属性信息产品系列包括关于地理区域特征的数据(例如人口和住房统计、陆地面积和人口密度)以及地理区域之间的关系。这些产品可独立于空间信息产品使用，也可以与空间信息产品一起使用(如果了解更多信息，请联系 Joe Kresovic: Joe.Kresovic@statcan.gc.ca)。

## F. 概要和结论

6. 188. 第六章研究了与点查后的地理任务有关的问题，包括数据的发布。在收到来自实地的更改地理数据库的报告后，一个程序便启动了，通过该程序持续对地理数据库进行修订，以便在调查和将来的人口普查中使用。数据的发布可以是在修订后的点查区层面或者通过并入新的发布单位，例如集群。根据人口普查可以生成大量静态或动态的产品和服务，包括地图、报告、地图集、CD-ROM和互动式网站。

6. 189. 在分解的层面发布数据带来了许多披露和保密的问题。另一个与数据发布有关的问题是免费提供数据，还是通过出售附加值产品(例如数据CD ROM或DVD)创收。本章还包括对地理数据产品的说明，例如地图阅读器、用于商业地理信息系统软件包的属性化空间文件以及因特网绘图产品。

## 书目和参考资料

- Aronoff, S. (1991年)。《地理信息系统：管理前景》。渥太华：WDL Publications。
- ASCE (1994年)。《制图学术语表》。Bethesda, 马里兰州：美国照相测量法和遥感技术学会，以及美国土木工程师学会。
- Bertin, J. (1983年)。《图像符号学：图形、网络、地图》。威斯康星州麦迪逊：威斯康星大学出版社。《图像符号学》，原文为法文，1977年，巴黎。
- Brewer, C. (1994年)。《制图和直观化的颜色使用准则》，摘自《现代制图的直观化》，A. M. MacEachren和D. R. F. Taylor编辑。伦敦：珀盖出版社。
- \_\_\_\_\_ (2005年)。《设计更好的地图：地理信息系统用户指南》。Redlands, 加利福尼亚州：ESRI Press。
- Bugayevskiy, L. M. 和J. P. Snyder (1995年)。《地图投影：参考手册》。伦敦，泰勒与弗朗西斯公司。
- Canter, F. 和H. Declair (1989年)。《世界展望：世界地图投影详细目录》。纽约：约翰·威利父子公司。
- Castro, M. 及其他人 (2004年)。《综合的城市疟疾控制：坦桑尼亚达累斯萨拉姆的一个案例研究》，《美国医疗卫生杂志》，第103-117页。
- Dent, B. D. (1999年)。《制图学：专题地图设计》，第五版。爱奥瓦州杜布克：Wm. C. Brown出版社。
- Duke-Williams, O. 和Rees, P. H. (1998年)。《普查办公室能否为多个小地区出版统计资料？对统计公布中误差问题的分析》，《国际地理信息学学报》，第12卷，第6期。
- 非洲经济委员会 (2007年)。《为非洲确定基本的数据集：社会经济发展中的地理信息》。文件ECA/ISTD/GEO/2007/02E。亚的斯亚贝巴。
- Fothergill, S. 和J. Vincent (1985年)。《国家状态：八十年代英国地图集》。伦敦：潘神图书公司。
- Garson, G. D. (2003年)。《公共信息技术：政策和管理问题》。宾夕法尼亚州赫希：思想出版集团。

- Hohl, P. 编辑(1998年)。《地理信息系统数据转换：战略、技术、管理》。新墨西哥州圣菲市：Onword出版社。
- Hurskainen, P. 和P. Pellikka(2006年)。《利用不同时间的航摄照片探查非正规规定居点的变化：肯尼亚东南部Voi的例子》。非洲环境遥感协会第五次会议的会议记录，赫尔辛基。
- Jensen, J. R.(1996年)。《数字图像处理简介：遥感前景》，第2版。纽约：普兰蒂斯·霍尔公司。
- \_\_\_\_\_和D. Cowen(1999年)。《城市/郊区基础设施和社会经济特征的遥感》，《摄影测量工程和遥感》，第65卷，第5期，第611-622页。
- Johnson, J. 和H. J. Onsrud(1995年)。《费用回收值不值？》，《城市与地区情报系统协会年度会议论文集》，7月，得克萨斯州圣安东尼奥。
- Khosrow-Pour, M. (2005年)。《实践电子政务：全球观》，宾夕法尼亚州赫希：思想出版集团。
- Kraak, M. J. 和F. J. Ormeling(1997年)。《制图——空间数据的直观化》。联合王国埃塞克斯郡哈洛：朗文集团公司。
- Krygier, J. 和D. Wood(2005年)。《制作地图：地理信息系统地图设计直观指南》。纽约：Guilford Pres。
- Longley, P., M. Goodchild, D. Maguire和D. Rhind。(2005年)。《地理信息系统和科学》。纽约：约翰·威利父子公司。
- Lynch, M. 和K. E. Foote(1997年)。《与地理信息系统有关的法律问题：地理学家的工艺项目》。奥斯汀：得克萨斯大学。
- MacEachren, A. M.(1994年)。《地图原理：符号化与设计入门》。哥伦比亚特区华盛顿：美国地理学家协会。
- \_\_\_\_\_ (1995年)。《地图如何起作用：代表性、直观化与设计》。纽约：Guilford出版社。
- McDonnell, R. 和K. Kemp(1995年)。《国际地理信息系统字典》。联合王国剑桥：国际地理信息公司。
- Monmonier, M.(1993年)。《详细制图：人文和社会科学的说明性制图》。芝加哥：芝加哥大学出版社。
- Montana, L. 和J. Spencer(2004年)。《将地理信息纳入测量调查：全球定位系统数据收集实地指南》。Chapel Hill, 北卡罗莱纳州：北卡罗莱纳大学卡罗莱纳人口中心。
- 荷兰综合学科人口研究所(1996年)。《人口普查和大规模人口统计创新技术专家小组会议论文集》，1996年4月22日-26日。海牙：荷兰综合学科人口研究所和联合国人口基金。

- Niebert, D. D. 编辑(2004年)。《开发空间数据基础结构：空间数据基础结构详细说明书》。全球空间数据基础结构。
- O’Looney, J. A. (2002年)。《用线连接政府：公共管理人员的挑战和机遇》。康涅狄格州西港：Quorum Books。
- Padmanabhan, G., J. Yoon和M. Leipnik (1992年)。《地理信息系统术语汇编》，《技术报告》，第92-13号。加利福尼亚州圣巴巴拉：国家地理信息和分析中心。
- Prévost, Y. 和P. Gilruth (1997年)。《撒哈拉以南非洲地区的环境信息系统》，2025年非洲的结构单元，第12号论文。哥伦比亚特区华盛顿世界银行和纽约开发计划署和联合国统计处。
- Rhind, D. 编辑(1997年)。《世界框架》。联合王国剑桥：国际地理信息公司。
- Robinson, A. H., J. L. Morrison, P. C. Muehrcke, A. J. Kimerling和S. C. Guptill (1995年)。《制图要素》第6版。纽约：约翰·威利父子公司。
- Snyder, J. P. (1982年)。《美国地质测量使用的地图投影》。哥伦比亚特区华盛顿：政府印刷局。
- \_\_\_\_\_ (1993年)。《使地球变平：两千年的地图投影》。芝加哥：芝加哥大学出版社。
- Tomlinson, R. (2007年)。《考虑地理信息系统：主管们对地理信息系统的规划》。Redlands, 加利福尼亚州：ESRI Press。
- Tufte, E. R. (1983年)。《定量信息的直观显示》。康涅狄格州柴郡：图像出版社。
- \_\_\_\_\_ (1990年)。《想象信息》。康涅狄格州柴郡：图像出版社。
- 联合国 (2007年)。《用于人口统计的地理信息系统》。出售品编号：E.97.XVII.30。
- \_\_\_\_\_ (2008年)。《人口和住房普查原则和建议，订正版第2版》，联合国出版物，出售品编号：E.07.XVII.8。
- 美国人口普查局 (1978年)。《人口普查和调查的地图绘制》，统计培训文件ISP-TR-3。哥伦比亚特区华盛顿：美国商业部，人口普查局。
- 美国联邦地理数据委员会 (1997年a)。《框架介绍和指南》。哥伦比亚特区华盛顿。
- \_\_\_\_\_ (1997年b)文化和人口数据小组委员会。哥伦比亚特区华盛顿 ([www.census.gov/geo/www/standards/scdd/index.html](http://www.census.gov/geo/www/standards/scdd/index.html))。
- 美国国家健康统计中心 (1997年)。《美国死亡率地图集》。哥伦比亚特区华盛顿：疾病控制和预防中心。

美国国家研究委员会(2007年)。《估计自然灾害和复杂的人道主义危机的高危人群的手段和方法》。哥伦比亚特区华盛顿: National Academy Press。

Weeks, J.(2007年)。《我们能从空中发现一个居民区吗? 界定加纳阿克拉的居民区结构》。GeoJournal, 69:9-22。

White, J. D.(2007年)。《在公共部门管理信息》。Armonk, 纽约: M.E. Sharpe。

Yankson, P. 及其他人(2004年)。《在遥感图像的基础上测量城市植被的灵活模式》。中国: 武汉大学。

## 附件一

# 地理信息系统

### A. 地理信息系统概述

A1.1. 地理信息系统指一套基于计算机的工具，用于输入、存储、管理、检索、更新、分析和输出信息。地理信息系统的信息与地理位置或区域的特征有关。换言之，地理信息系统让我们回答事物在何处或在某一特定位置有什么事物的问题。

A1.2. “地理信息系统”一词在不同的语境中有不同的含义，既可指处理空间信息的整套软硬件，也可以指专用于处理地理特征信息的具体软件包，还可以指一个国家或地区在地理方面的应用程序，如一套综合性地理数据库。最后，它有时还可能用来说明一种涉及地理空间数据处理、计算和方法的研究领域。例如，有些大学已经开设了地理信息系统学位课程。人们越来越多地用“地理信息学”来指与基于计算机的地理软件和程序有关的科学研究。与此相关的一个术语是地理空间技术。这个术语的范围更广泛，包含地理信息系统、遥感和全球定位系统技术。

A1.3. 很多领域对地理信息系统的建立均做出了贡献，如图A1.1所示。测绘行业对测量和表达真实世界的地理特征提供了多种尺度和工具；计算机科学提供了地理信息存储及管理的框架结构，并和数学一起提供了处理表达真实世界地理特征的几何对象的各种工具。通过社会经济、环境和地形调查得出的大量数据，地理信息系统可在各种领域得到广泛应用，如考古学或海洋学等以学术为主的领域，乃至包括市场营销或房地产在内的商务领域。

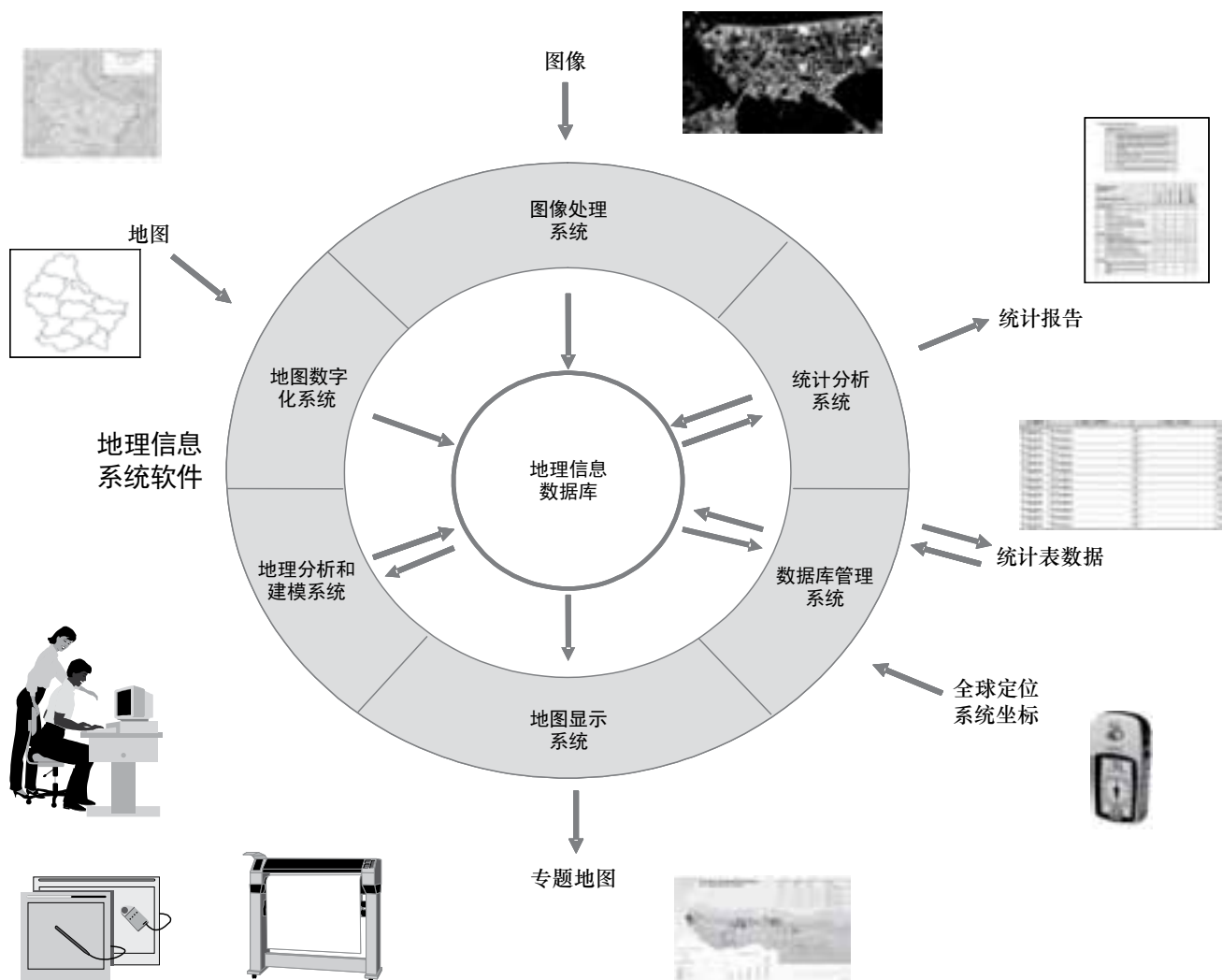
A1.4. 产品目录类型的应用广泛见于公用设施部门，如电话公司通过地理数据库对其基础设施进行管理和维护。另一个实例是地区和地方政府部门使用的土地命名系统。在其他一些领域，地理信息系统用于支持数据收集。将数字化地图制作和地理数据库用于人口普查和数据发布无异是本《手册》中最为恰当和典型的例证。大多数分析性应用见于科研部门以及很多实用领域，例如自然资源管理或市场营销。森林开发公司通过地理信息系统尽可能实现林木采伐的可持续性。营销公司或零售公司也利用复杂的空间分析瞄准客户或为新店面选址。

#### 1. 硬件、软件和数据

A1.5. 硬件和软件问题在本《手册》关于用于人口普查活动的地理数据库的建立和维护的部分已经讨论过。一般来说，与用于其他面向图形的、以处



图A1.1  
地理信息系统的基础



资料来源：Elbener/世界卫生组织。

理大量数据为特点的应用相比，地理信息系统所需的硬件基本相同：高端PC兼容计算机或工作站、一台大型高分辨率显示器和通用的输入装置——键盘和鼠标。大版面的数字化板或扫描仪用于把纸张印刷的地图转换成数字化数据库。建筑师和图形设计师也是使用这类工具。大版面绘图仪和台式打印机用于地图输出，以供显示和直观分析。

A1.6. 地理信息系统软件近年来发展迅速，从难以掌握的命令行式的系统发展到便于使用的菜单驱动软件包，任何人只要稍加培训即可使用。分析人员使用高端软件包来创建新型地理数据库并执行高级的空间数据分析。对于中档软件而言，目前有多种桌面地图绘制软件，它们不但有标准界面，而且在数据输入、管理以及分析和输出方面具有广泛的能力。最后，对于低端软件来说，可以选用地理数据浏览器。用户在使用这种软件时虽不能改动数据，但却拥有多种显示功能。这类软件有些是免费发布的，属于性能卓越的数据传播工具。

A1.7. 最近的一项进展是几家地理信息系统销售商出售的与地理信息系统有关的软件脚本或例程序工具箱。用户利用它可以在具有行业标准的、面向对象的编程环境中根据自己的需要定做应用软件。它既可以是一套独立的系统，也可以集成到其他的软件包中。其中有些产品还包括用于开发基于因特网的地图绘制应用软件所需的工具。

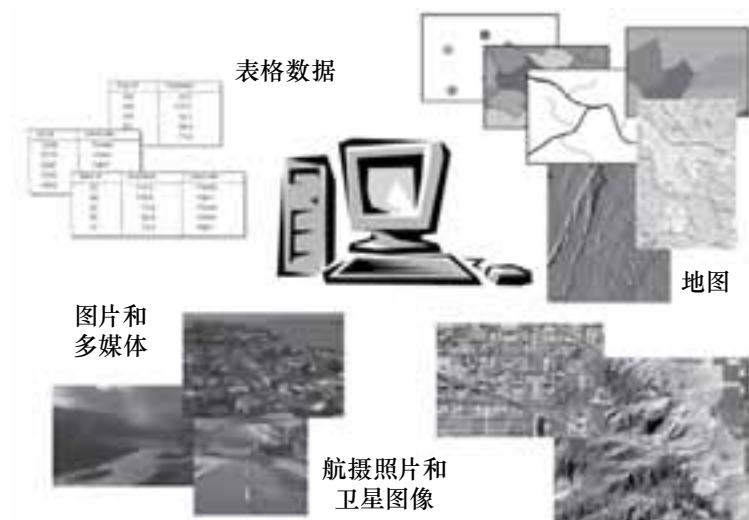
A1.8. 当前，地理信息系统软件的发展趋势集中在两个方面：因特网地图绘制和模块设计。模块设计可将地理空间功能集成到任何应用软件中。用户通过自己的网络浏览器和根据命令下载的软件，就能在远程地理空间数据库上进行数据的查询和分析。在高端应用方面，还把地理信息系统与关系型数据库管理系统结合在一起。正如地理信息系统软件通过关系型数据库管理系统存储和处理属性信息那样，某些数据库管理系统也已经包括一些存储和处理地理对象的功能。地理信息系统与其他信息系统之间的区别正在消失。

A1.9. 数据是供地理信息系统应用软件使用的燃料(见图A1.2)。很多最普通的地理信息系统数据集是纸张地图的数字化对等物，如以不关联层显示道路、河流、标高以及居民点情况的地形图。专题信息包括以行政单位为基准的社会经济属性、展示植被或土地利用状况的说明性地图、汇水或流域边界等派生指数。数字化地图上的任何地理目标均可通过与数字化空间数据库连接在一起的数据表加以说明。有时，只需少量的属性信息就可以说明一组地理特征。而在其他情况下，例如人口普查数据库，系统存储的属性信息可能十分丰富。

A1.10. 遥感(即来源于航空拍摄或卫星的)图像是地理信息的另一来源。低空飞行的飞机拍摄的照片以及卫星发回的图像也可与其他空间基准信息集成在一起。有时，这些图像仅作为背景用于专题地图或地形图，但更常见的情况

图A1.2

存储在地理信息系统内的各种类型的信息



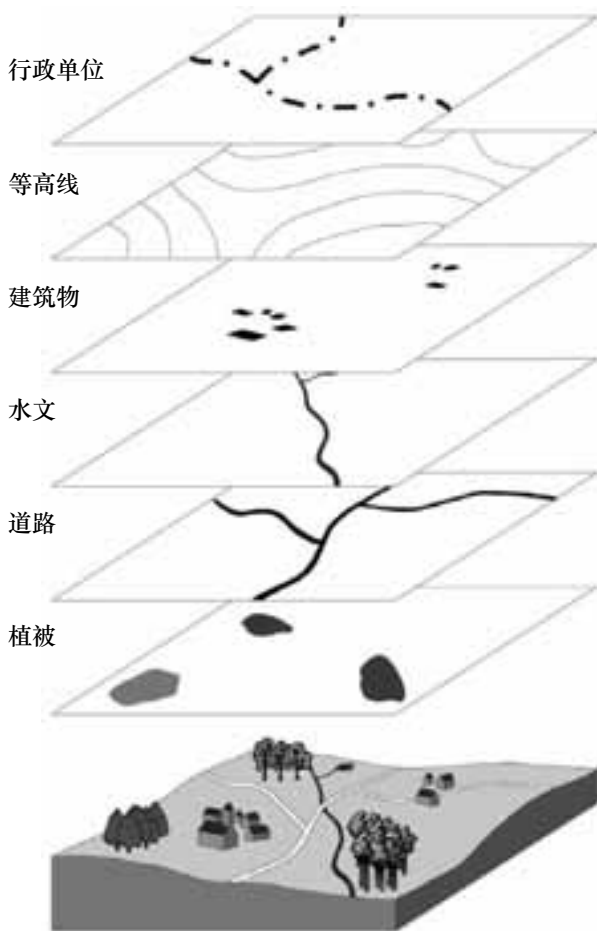
是，解释和提取图像中所包含的信息，然后作为数字化地图信息存储起来。最后是多媒体信息，如照片、录像、文本乃至声音等均可集成到地理信息系统内。通常，这种集成通过热连接的手段实现。用户利用交互方式通过点击某一个地理特征即可查看有关地理位置的照片或录像。易于使用的因特网制图程序（例如Google Earth）的出现对此提供了很大的方便。

## 2. 地理数据层

A1.11. 地理信息系统数据库是对真实世界的一种基于计算机的表达方式。地理信息系统软件提供的是对空间界定的地理特征信息进行组织的工具，而数据层是地理信息系统的基本组织原则，如图A1.3所示。和地形图不同的是，地理信息系统没有把所有的空间特征存储在同一个地方，而是把类似特征按组结合在一起，存储在一系列数据层中。

A1.12. 地理数据库可以包括由实际地理特征组成的层，如道路、河流和建筑等，以及一些在地面上看不到的具有界定特征的数据层，如行政区边界和

图A1.3  
数据层——作为索引系统的空间



邮政区域等。另外，地理信息系统软件允许用户在现有的数据层上新建一些数据层。例如，一个新的数据层可以显示由数字化标高线数据推导出来的流域范围，或者某个医院四周规定距离内的所有街区。

A1.13. 在建立多层地理数据库的过程中，地理特征可以从一系列不同的地形图和专题地图中提取。另外，实地观测数据以及卫星或航摄照片提供的遥感数据通常与地图数据集成在一起。地理信息系统提供的工具把所有不同数据集成在由地理坐标系统定义的通用基准框架内，使用户能把不同类型的的数据组合到一起，根据几个数据层的内容创建新信息，或进行涉及多个数据层数据的复杂查询。这种以地理位置作为链接，集成各种不同来源数据的能力有时称为“以空间作为索引的系统”。这的确是地理信息系统最重要的优点之一。

## B. 地理信息系统数据模式

A1.14. 尽管各种各样的信息都可以存入地理数据库，然而只有少量几种空间信息的表达方法。为了开发地理信息系统应用软件，真实世界的特征值要转换成便于计算机存储和处理的简化表达形式。对此，有两种数据模式——即内部数字化表达方式——是商业地理信息系统软件采用的主流模式，尽管许多程序如今可以轻易地处理两种数据模式。“向量”数据模式用于把不连续的特征符号化，例如房屋、道路或区域等。还有一种是“光栅”数据模式，通常用来表达连续改变的现象，如标高或气候，此外还可用于存储卫星图像或航摄照片。尽管很多辅助性的数据更适合用光栅模式存储，对人口普查应用来说，向量数据模式更为有用。

### 1. 向量

A1.15. 向量地理信息系统利用一系列所谓的几何元素来表达真实世界的特征：点、线和多边形(见图A1.4)。在计算机数据库中，一个点用x、y坐标表达。一条线是一系列顺序排列的x、y坐标，其中，端点通常称为节点，而中间的各点则称为顶点。多边形或面积则用一系列封闭的线段表示，在这个回路中

图A1.4  
点、线和多边形



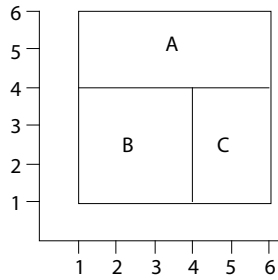
第一个点与最后一个是等同的。点可以用来表示房屋、水井或测量控制点；线条可以代表道路、河流一类的特征；而点查区或区域等则以多边形表示。

A1.16. 最简单的存储数据的向量数据模式不必在地理特征之间建立任何联系(见图A1.5)，这种模式有时被称为“意大利面条模式”(例如见Aronoff, 1991年)，因为线条在数据库内重叠而不相交，像面条一样堆在碗里。更复杂的“拓扑式数据模式”可存储数据库内不同特征之间的关系。例如，将交叉的线条分开并在相交处增添一个节点。为了避免邻接的多边形边界出现两次定义的情况(每一封闭的多边形各定义一次)，因此规定线段只能存储一次，同时把多边形分别处于线的左侧还是右侧这样的信息一起存储起来。关于节点、线和多边形之间的关系信息存储在属性表内。

A1.17. 只要我们看一看空间数据库能够解决什么问题，拓扑式数据结构的优点就十分明显了。一个拓扑型结构的空间数据库能够进行快速查询，查找单独的数据对象以及它们与其他数据对象的关系。例如，迅速确定一个特定点查区的所有邻接区，此时，系统将找出界定该点查区的线段列表，并找出与这些线段连在一起的其余的所有点查区。

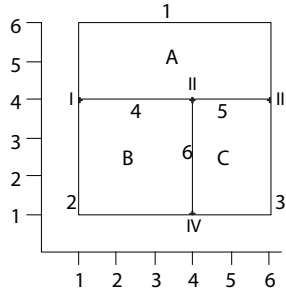
图A1.5  
向量数据模式：面条模型与拓扑模式的对照

“意大利面条式”数据结构



多边形	坐标
A	(1,4), (1,6), (6,6), (6,4), (4,4), (1,4)
B	(1,4), (4,4), (4,1), (1,1), (1,4)
C	(4,4), (6,4), (6,1), (4,1), (4,4)

拓扑式数据结构



O = 外部多边形

节点	X	Y	线
I	1	4	1,2,4
II	4	4	4,5,6
III	6	4	1,3,5
IV	4	1	2,3,6

多边形	线
A	1,4,5
B	2,4,6
C	3,5,6

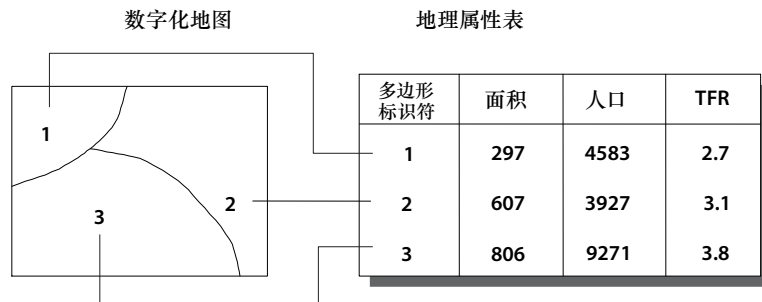
线	来自节点	至节点	左侧多边形	右侧多边形
1	I	III	O	A
2	I	IV	B	O
3	III	IV	O	C
4	I	II	A	B
5	II	III	A	C
6	II	IV	C	B

A1.18. 如今大多数地理信息系统软件包采用全拓扑型数据结构，使得诸如多边形重叠一类的复杂操作成为可能。在重叠操作中，要把两个向量数据集，如行政区和流域边界，组合到一起。由于两方面输入的数据集出现相交，因此生成一些新的较小的多边形。大多数桌面地图绘制系统使用比较简单的数据结构，其中，所有多边形定义为封闭回路，这样，定义两个区间邻界的线段就在数据库中存储了两次。

A1.19. 数据库内的每一个特征均被赋予了一个惟一的标识符，藉此把几何特征与数据或属性表内的相关条目链接在一起(见图A1.6)。用户可以把特征信息增添到相应的数据库记录上。对于代表房屋的点，用户可以列出通信地址、房屋类型以及是否配备电气和卫生设施的清单。在点查区数据库内，用户可以增添官方行政代码、居住单元数量以及任何为该点查区编译的人口普查数据。出于实用目的，大多数地理信息系统均采用关系型数据库模式，分别存储属性和非空间信息。属性文件与数字化地理数据紧密地集成在一起，并通过地理信息系统或关系型数据库管理系统进行访问。

图A1.6

向量型地理信息系统中存储的空间和非空间数据

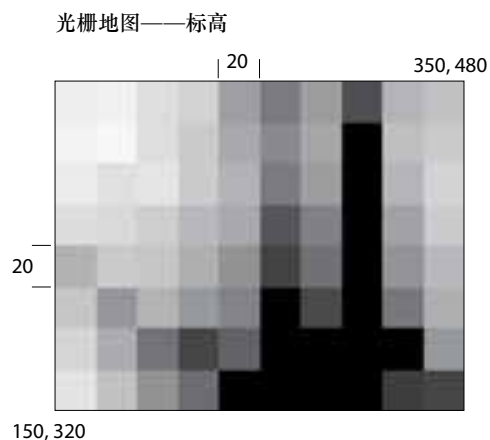


A1.20. 有些桌面地图绘制软件在这两个极端情况之间，即在最简单的“意大利面条模式”和最复杂的全拓扑模式之间达成了妥协。它们虽非全拓扑型的，但仍然可快速计算邻接和连接信息。它们综合了简单数据模式容易编辑的特点与拓扑向量地理信息系统数据模式分析能力强大的特点。

## 2. 光栅

A1.21. 光栅地理信息系统软件包把空间分成行列相交的规则阵列。阵列或网格中的一个单元格有时称为像素，即图像的元素，并显示出该数据模型在遥感和图像处理中的原始状况。在大多数的光栅系统中，某一指定位置的属性值(如标高)就存储在光栅内的相应单元格中。因此，标高光栅数据库干脆就是一长串的标高值。系统需要的附加信息只是光栅图像的行数和列数、用真实世界单位(例如米或英尺)表示的单元格(通常是一个方块)的大小，以及整个光栅的某一角的坐标值(见图A1.7)。此类信息通常存储在数据结构的文件头内，或

图A1.7  
光栅数据文件实例



ASCII光栅文件									
列数	10								
行数	8								
左下角 x	150								
左下角 y	320								
单元格尺寸 (方格)	20								
219	313	407	462	681	783	689	877	595	540
297	274	407	501	642	744	650	955	556	501
336	391	368	501	603	783	689	994	595	462
414	430	485	579	642	861	767	1072	673	501
609	508	524	618	720	900	806	1267	712	579
531	703	602	696	759	978	884	1189	790	618
453	625	797	891	837	1173	1079	1111	985	696
375	547	719	813	1032	1095	1001	1033	907	891

另存成一个小文件。通过这些信息，系统可以计算出网格的尺寸。例如，右上角x坐标为 $150+10 \times 20=350$ 。系统利用这些信息就可以把其他地理数据层与光栅网格正确地登记在一起，例如，从网格顶部提取向量特征。

A1.22. 如果光栅中有很多相同的数值，那么这种数据存储方法的效率则非常低。例如，离散对象有时也存成光栅格式。光栅格式的地区图将在每一个单元格中显示该地区的标识符，或单元格内的总人口数。很明显，很多连续的单元格中经常含有同样的数值。因此，大多数的光栅式地理信息系统均采用某种数据压缩形式。其中最简单的就是长线程编码，系统在这种方式下仅存储两个成对的数值：数据值和该数据值所重复的次数，从而大大减小了文件的尺寸。

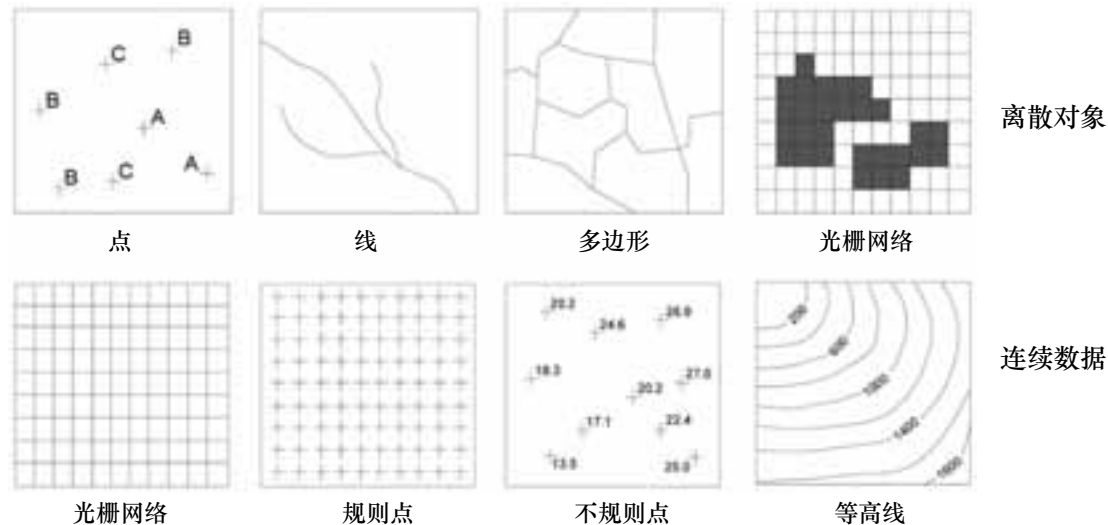
A1.23. 光栅数据经常用于存储连续多变的数据，或具有多种连续灰色调的图像。正如离散对象可以通过光栅格式显示那样，连续的数据也可以用向量数据结构表示。最好的例子是等高线，它在地形图中表示标高。其他的例子见图A1.8。

### 3. 向量和光栅数据模式的优缺点

A1.24. 光栅数据模式的长处在于简单。在光栅式地理信息系统内，很多地理数据的操作比较容易并且执行速度较快。与标高或水文数据一样，模式化的连续数据都是经常通过光栅式地理信息系统实现的。光栅数据模式的缺点在于，光栅数据集的大小与所表达的空间特征的精度之间常常无法兼顾。用非常致密的光栅网格可以细致入微地表示边界上的所有曲线，但需要极大的磁盘空间。就利用“对象”（例如点、线和多边形）进行地理分析而言，向量模式更好用。

图A1.8

向量和光栅数据模式均可用于显示离散和连续数据



A1.25. 大多数的地理信息系统可以按照两种数据模式操作。采用哪一种数据模式更好取决于应用环境。对于人口普查和很多其他社会经济应用来说，更宜采用向量模式。向量数据结构可以更紧凑地表示界定社会经济对象的点和多边形。它与数据库管理系统之间的紧密联系，能够支持以大量属性信息为特点的社会经济方面的应用。例如，人口普查或调查经常涉及成百上千个变量，它们与一定数量的空间特征(如人口普查区、村庄或调查集群等)密切相关。最后，向量式地理信息系统数据库打印出的成品更接近传统技术印制的地图。

A1.26. 即使如此，光栅数据的处理能力在人口应用方面的重要性与日俱增。某些用于给点查区划界的输入数据就是以光栅形式输入的边界，例如，本《手册》第四章讨论过的通过遥感图像建立或更新人口普查地图的工作就是如此。不过，两种数据模式的选择通常不是“非此即彼”。现在，很多地理信息系统软件包均支持两种类型的空间数据。例如，可以用光栅数据作为背景，在上面绘制线条或多边形特征。因此，遥感图像或标高表面与相关信息可以同时显示在计算机屏幕上，以便确定点查区的轮廓。

#### 4. 精确度和准确度

A1.27. 精确度和准确度这两个词虽然经常互换使用，但仍有不同的含义。例如，在空间方面，准确度指的是所表示的位置与地球表面的真实位置的相符程度。而精确度指的是，在测量过程中，区分两个微小数量或微小距离的能力。例如，如果我们的测绘工具只能以米为单位测量坐标，那么地理信息系统中的点位置最多只能归入相近的以米为单位的单元格中。如果有更精确的测量工具，则所得点坐标可以精确到厘米或毫米。

A1.28. 实际上，向量地理信息系统的坐标精确度可以说是无限的，因为它们可以使用双精度数据类型存储地理坐标（此处的双精度系指用8个字节表示



一个浮点数)。而空间坐标的准确性大部分取决于数据采集工具。工程上或研究大地板块构造时使用的测绘工具,最高准确度可达到1毫米。然而,地理信息系统使用的大多数数据,由于数据来源的原因,准确度都不会太高。比如,有的数据来自纸张式地图,有的来自手持式全球定位系统,有的甚至来自野外作业时使用的草图。这种情况下,其准确度只能以米来计算,而不是毫米。

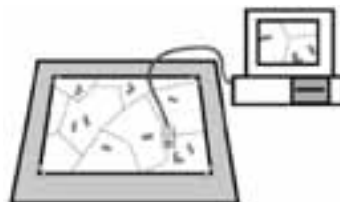
### C. 地理信息系统的功能

A1.29. 下面的表格概括了地理信息系统工作能力的情况。这个表格很不完备,因为高端地理信息系统软件、甚至桌面地图绘制软件还有很多用于数据输入、处理、分析以及显示的专用功能。

#### 数据输入和管理

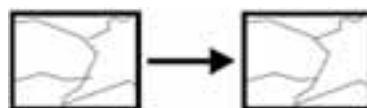
##### 线条跟踪, 坐标数据输入

目前最常用的仍然是借助于数码化表的坐标数据输入方式。用鼠标在纸制的地图上跟踪线条并将捕获的数据存入地理信息系统或数字化软件。此外,还可对地图进行扫描,建立光栅位图,然后转换为向量格式。



##### 编辑

对线条进行数字化处理以后,对数据进行检查,排除错误。常见的问题有,线条连接不良(不到位或过头)、线条丢失、重复处理等。其中的一些操作可在地理信息系统中自动进行。



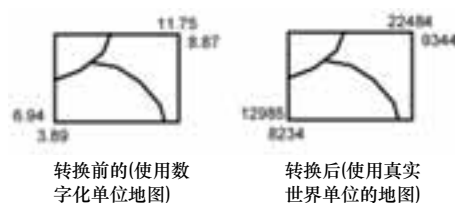
##### 建立拓扑结构

经过数字化或向量化处理的线条之间没有任何联系。地理信息系统软件可以计算数据集内特征之间的邻接和连接关系。



##### 地理基准与投影变化

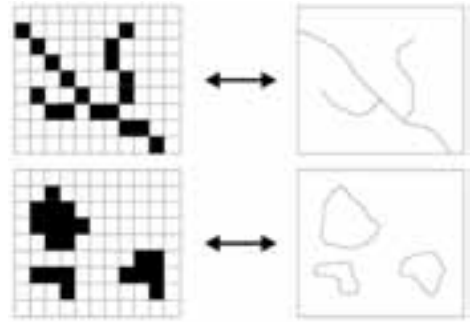
数字化的线条以厘米或英寸为单位,需要转换成真实世界使用的、与原地图坐标系单位相对应的单位,例如米或英尺。为了进行数据集成,数字化地图的投影方法也需要做相应的变化。



光栅与向量之间的转换

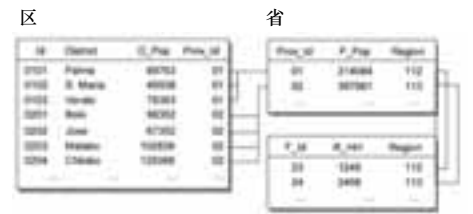
大多数商业地理信息系统软件均支持某种格式的光栅图像。因为每一种数据模式适合于不同的任务，所以应该具有将一种模式转换为另一种模式的功能。光栅到向量的转换功能也用于自动转换扫描的地图。

相反的操作，即向量到光栅的转换，在光栅式地理信息系统进行分析和模式化操作时需要使用。



属性数据管理

数据库内的每一个特征都有一个唯一的标识符，用于链接有关地理特征的外部信息。为了能处理和分析属性表格，地理信息系统通常与一个关系型数据库管理系统集成在一起。



重新分类、加总

地理信息系统允许根据通用标识符对特征进行加总。例如，对几个点查区进行分组，构成人口大致相等的普查作业区。



子集的建立、切块处理

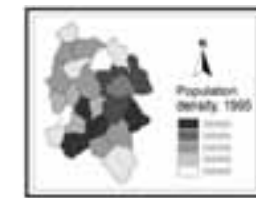
除了根据查询选择子集之外，地理信息系统也可以利用所谓的切块操作建立特定的子集。



显示

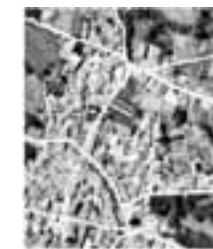
地图绘制功能

制作地图输出展示只是地理信息系统地图绘制功能的一种应用。地图绘制中的符号处理，对于在屏编辑和分析中不同特征的区分也十分重要。

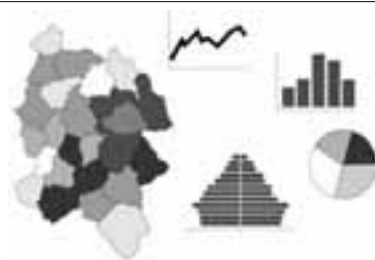


图像与向量数据的合并显示

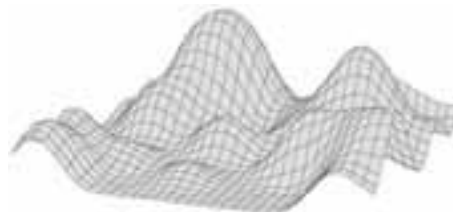
图像或光栅数据有多种不同的来源：扫描的地图、遥感图像以及光栅式地理信息系统数据，它们都以某种网格形式加以存储。同时显示向量和光栅数据可为分析提供有价值的工作环境，能够从光栅数据中有选择地提取特征。



**链接到统计图表** 对空间数据进行以数据为主的分析通常采用地图绘制与属性数据检查相结合的方法。统计图表特别有用，尤其是将其放到地图上时。



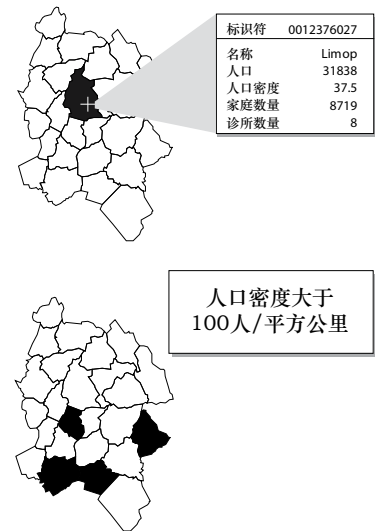
**各种表面的三维立体显示** 诸如标高或降雨量乃至人口密度等连续数据可以通过不同形式予以显示：光栅网格、等高线或通过网状框架模拟出的三维可视化图形，在网状框架上可展示其他地理特征。



**查询**

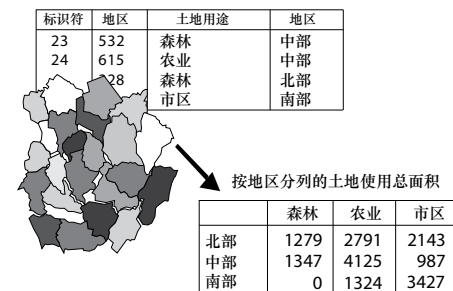
**空间数据库的查询**

“在……有什么？”和“何处是……？”这些问题是地理信息系统能够回答的最基本的地理问题。按照简单的浏览模式，用户可以在数字化地图上选择地理特征并获得有关信息。反之，用户也可以选择符合判别标准的地理特征，并将其显示在地图上。地理信息系统通常链接到数据库管理软件上，查询操作以SQL概念为基础。在地理信息系统内，也可根据诸如距离等地理关系进行查询(“在该地x公里之内是什么？”)或基于两个或多个地理信息系统数据层提问(“哪些建筑位于这个点查区？”)。



**属性汇总操作**

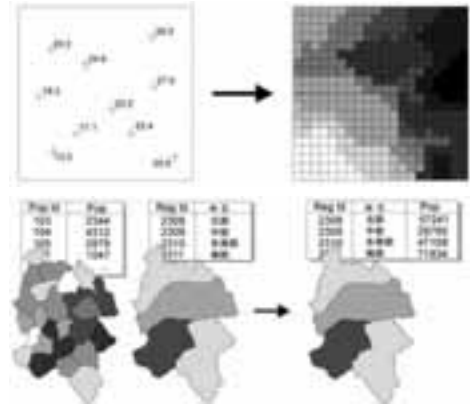
通过数据库操作，可以从地理信息系统数据集的地理属性表中提取出有用的统计数字总数或综合多个表格的数据。例如，计算表格内一个字段的的最小值、最大值以及平均值。或跨越表格中的两个或多个字段，生成每项综合属性第三个字段所需的总计数量。例如，还可以计算一个国家的某个地区之内，每一种土地用途的总面积。两个或多个地理信息系统层合并之后，通常需要通过多边形重叠操作进行跨表格的处理。



### 空间数据转换

#### 插值

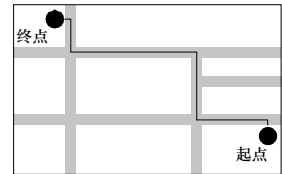
利用插值的方法可以根据样本数据建立一个完全覆盖区。例如，根据气象站的一套雨量记录表面建立显示整个地区降雨量的光栅表面。对于社会经济应用来说，更重要的是所谓的面积插值法。例如，运用某一地区的人口数估算出另一个边界不一致的环境监测区的人口数。



### 距离和时间操作

#### 简单的距离计算

计算距离是地理信息系统的基本操作之一。距离(和时间)可按直线计算，也可以按道路网络的距离计算。例如，根据地理信息系统道路数据库，可以估算出距离和旅行时间。



距离：3.2公里  
旅行时间：10分钟

#### 缓冲区

一种特殊的距离操作是建立缓冲区。缓冲区可以围绕着点、线或多边形建立，而且可以根据属性值进行加权。例如，铺设过的道路可以比土路得到更多的缓冲区。缓冲区常常用于空间查询。例如，为了确认河流两岸3公里范围内血吸虫病例的数量，在指定多边形内加入一个缓冲区，数据库查询便可依照次序执行。



#### 寻找最近的地理特征

如果需要确定特定类别内最近的地理特征，则综合运用数据库查询与距离计算的方法。例如，我们可能想要计算某一地区内所有地点到最近医院的距离。得出的地理信息系统数据集通常称为可访问性表面。



#### 泰森多边形

另一种类似“寻找最近的地理特征”的变体功能是把整个地区划分成若干个多边形，并指定最近的设施。得出的各面积单元称为泰森多边形。这项功能经常用于建立简单的汇水区或服务区。

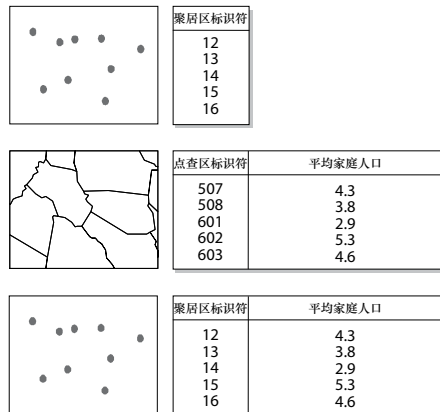


## 数据层的合并

### 多边形操作中的点线

很多地理信息系统能够帮助回答的问题都要求把几个数据集的信息合并在一起。例如，我们可以用一组点坐标代表人口普查时遇到的人口聚居区，把调查得到的信息与点查区得到的数据组合在一起。地理信息系统将确定每个点属于哪一个点查区，并将人口普查数据补充到调查点的属性记录中。

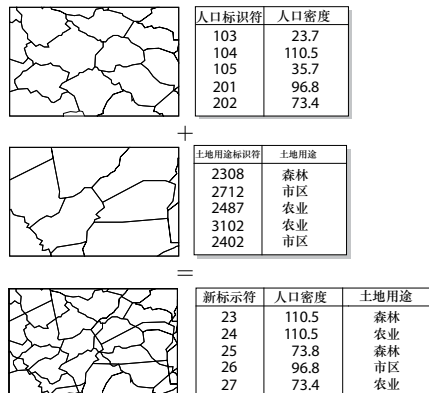
同样的操作可以汇总一组地区内点或线的特征属性。例如，可以利用家庭调查采样(点)确定每一个医疗保健区的平均生育率。



### 多边形重叠

把两个区域地理特征的地理信息系统数据集结合在一起称为多边形重叠。系统将对数据集进行合并，并根据重叠区建立起新的区域单位。得出的新数据集包含原有的两个数据集的属性。属性保持不变(如分类信息或比率)，还是划分给新的多边形(如计数数据)取决于数据类型。

多边形重叠操作常用于跨表格合并，例如，按土地用途计算人口普查数据等。



## 附件二

### 坐标系统和地图投影

#### A. 引言

A2.1. 附件一中对地理信息系统概念的评述强调了空间数据集成的好处。通过把不同类型的地理信息组织成数据层，可以利用来自很多不同作业区的数据进行测量、查询、模型化处理以及其他类型的分析。因此，人口普查数据可以与土地利用数据或农业生态数据组合在一起分析，或者，社会经济调查信息可以链接到基于患病风险的地理基准数据上。通过对不同数据层进行纵向集成，众多来源的数据相互链接的能力得以实现。这意味着所有地理数据集都以同一个坐标系为基准，不同数据层在顶部重叠时，可以正确地对准。

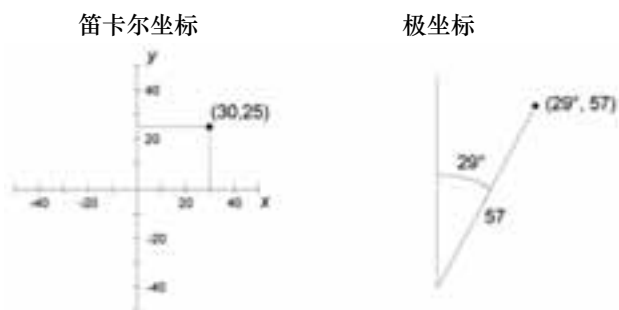
A2.2. 在建立地理数据库时，例如建立一个人口普查地理数据库，数据开发者必须保证从硬拷贝数据源、数字化地名字典或野外作业期间获得的空间坐标与边界数据是在一个所谓的“地理基准”过程中被标注在一个正确的坐标系中的。这也保证了那些为邻近区域分别开发的数字化地图同时显示在计算机屏幕或打印在页面上时，能够实现完美的匹配。

A2.3. 对于通过传统技术绘制的人口普查地图，我们无需考虑太多，因为这种纸质地图(常常是实地绘制的草图)仅仅用于点查目的而已。这种地图不与其他数据集成，也不在任何的空间分析中使用。过去，坐标系和地图投影的知识不如现在这么重要。现在建立的数字化数据库将用于不同目的。本附件将对重要的地图绘制概念进行一番简单的介绍。地图绘制的教科书，如Robinson和其他人(1995年)、Kraak和Ormeling(1996年)以及Dent(1998年)均提供了很多补充信息。对该课题更为专业化的探讨可参见Canters和Decleir(1989年)、Snyder(1993年)以及Bugayevskiy和Snyder(1995年)。

#### B. 坐标

A2.4. 在地图绘制过程中，测量物体在地球表面上的位置称为地理坐标系，有时也称为地理参照系。在二维几何中，最普通的坐标系即“笛卡尔坐标系”，根据法国科学家勒奈·笛卡尔(1596-1650年)命名。坐标给出的是两个固定数轴(x和y)上的，从固定原点进行计算的垂直距离。这就是地理信息系统使用的坐标系统，同时也更普遍地用于计算机图形软件中。还有一种定位方法是“极坐标系”，从一个固定的原点测量角度和距离(见图A2.1)。

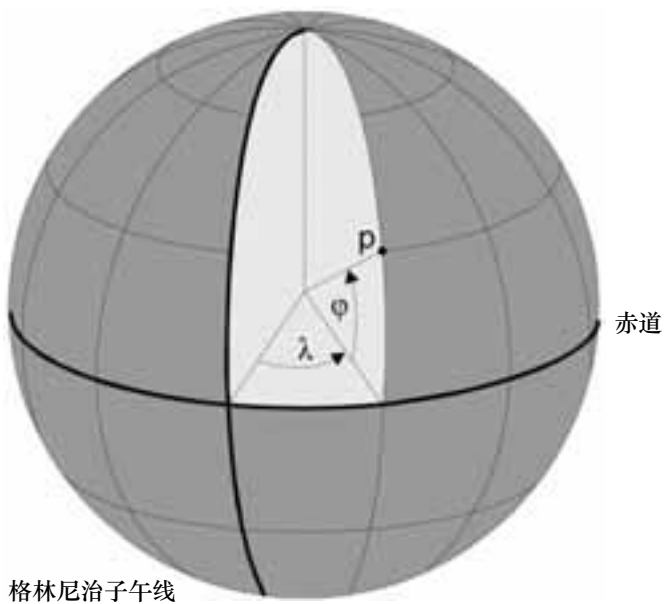
图A2.1  
平面坐标系和极坐标系



A2.5. 一幅平面地图，无论是在纸张上，还是在计算机屏幕上，总是把坐标显示在一个平面二维坐标系上，其坐标值按标准单位进行计算，如米或英尺。坐标通常以“x”和“y”命名，不过在地图绘制课本中，也常常使用“东向”和“北向”说法。然而，地图上的对象只是地球表面地理特征的一种表示方法。由于地球是球形的，所以地球表面的坐标是按照球面坐标系测定的。更为具体地说，通常是采用纬度和经度坐标进行基准定位，从而构成了一个球形的极坐标系。在这种坐标系中，任何一个点p的位置用两个角度加以界定，一个是纬度角 $\phi$ ，即相对于赤道平面的角度；另一个是经度角 $\lambda$ ，相对于0度平面或格林尼治子午线测定(见图A2.2)。

A2.6. 为了制作纸张印刷的世界或局域地图，这些球面的经纬度坐标需要按某种方式转换成平面坐标。在最近一本关于地图投影法的书中，作者把将三维地球的一部分表示成二维的过程称为“展平地球”(Snyder, 1993年)。

图A2.2  
球面坐标：纬度/经度基准系统

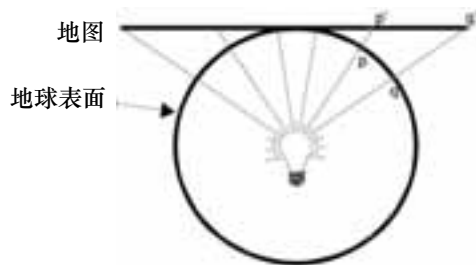


## 地图投影法

A2.7. 把地球经纬度坐标转换成平面坐标的数学过程叫做地图投影法。顾名思义，我们可以想象一下投影的过程。例如，想象有一个光源位于地球的中心，并且地球的表面是透明的，我们在上面勾画出了我们需要的地理特征。在地球顶部放一张纸，沿着投影出来的地形轮廓重新描一遍，描在这个所谓可展表面之上。例如，位于地球表面 $p$ 点的地理特征点将位于地图上的 $p'$ 点处。正如图A2.3所示，地图上与地球模型离开越远的点，其相对于切点距离的失真程度就越大。例如，自地球模型上 $p$ 和 $q$ 两个点之间的距离比地图上 $p'$ 和 $q'$ 两个点之间的距离要小得多。位于赤道上的点无法通过这种方法进行投影，因为通过赤道的光线与地图平行。因此，这种投影法只能用于相对靠近切点的区域。

图A2.3

地图投影过程(方位投影)图解



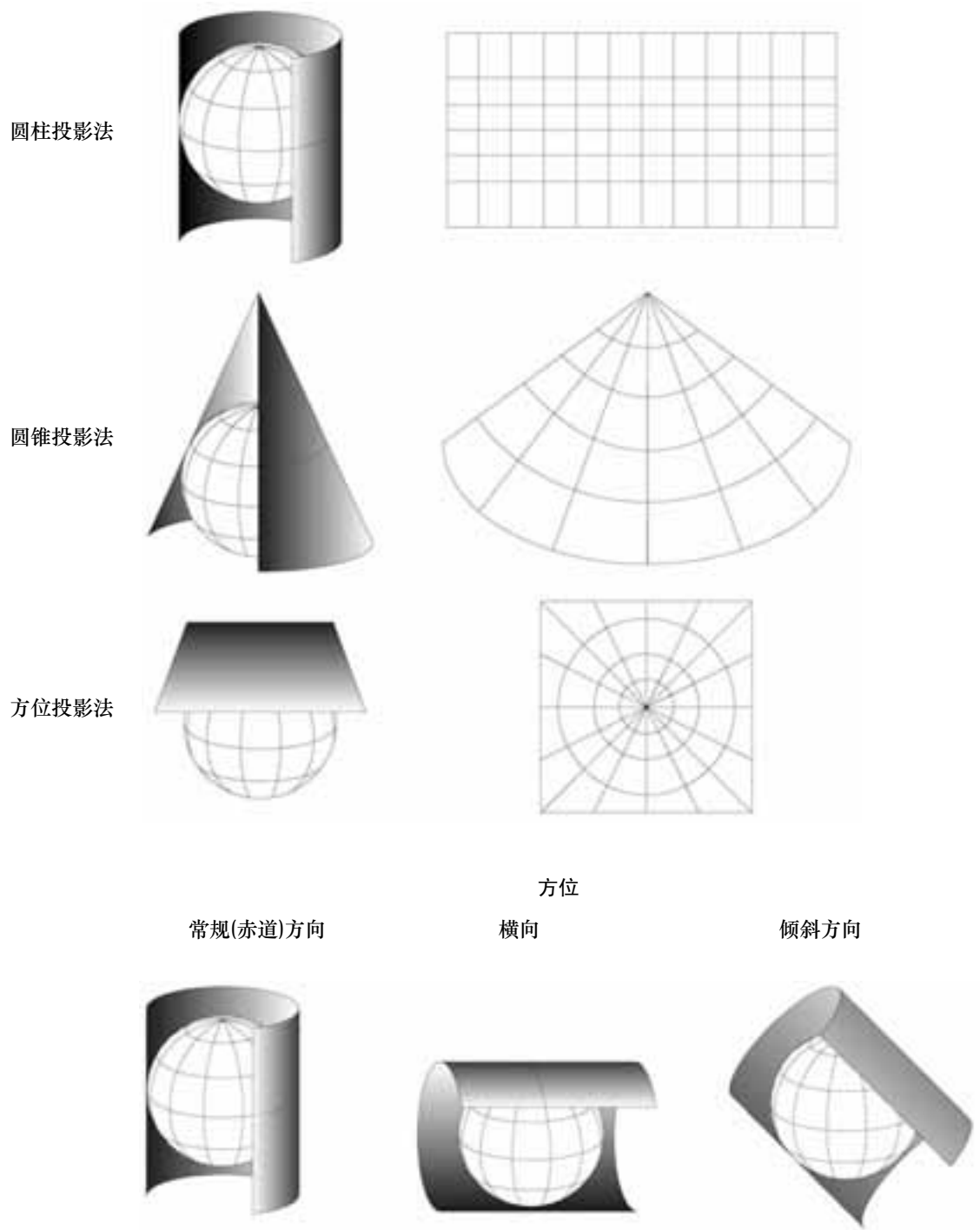
A2.8. 几百年来，地图绘制工作者开发出多种不同的地图投影法，这些方法可以按照地图的放置方式或地图包围地球模型的方式加以分类。图A2.4是一个概览，说明了三种地图投影法的构成。它们分别是圆柱投影法、圆锥投影法和方位投影法。如右侧地图经纬网所示，每种地图投影法都形成了一套独具特色的经纬度网格线。

A2.9. 地图绘制工作者也可以选择“可展表面”与球体的接触位置，不论是圆柱面、圆锥面还是平面均可以选择。接触线和接触点的周围通常是尺寸和形状失真最小的区域。如果要制作一幅世界某一特定地区的专用地图，则可以选择一下地图的投影“方位”，以使地图能以最佳状态进行表达。

A2.10. 假想的光源并不一定总是位于地球的中心(见图A2.5a.)，也可放在远端的极点上(见图A2.5b.)。此外，我们还可以想象光源不是点光源，而是群体的面光源，它们位于与地图平行的平面基底上，一起发出投射光线(见图A2.5c.)。按地图绘制行业的术语来说，这三种投射方式分别叫做：球心投影、球面投影和正射投影。从地图上 $p'$ 和 $q'$ 两个投影点可以看出，每一种假设，都会导致地图上的对应位置出现不同的失真。

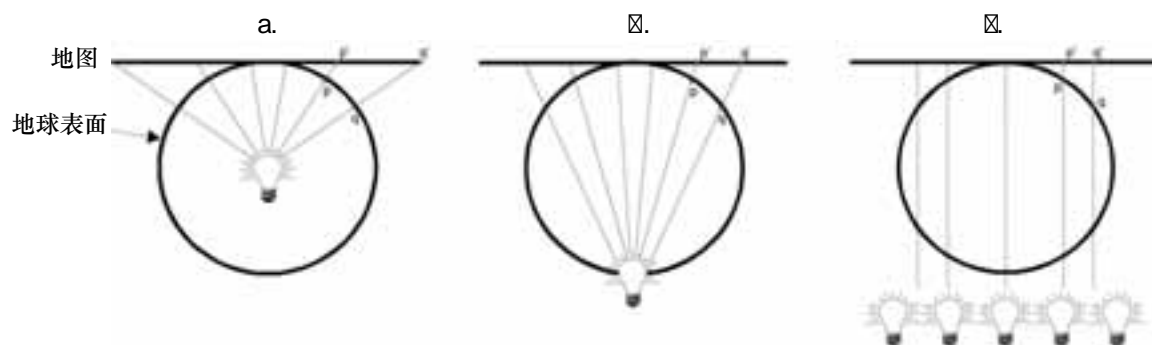


图A2.4  
各种地图投影法



图A2.5

## 构成投影的各种方式



## C. 地图投影的特性

A2. 11. 尽管想象的光源是说明地图投影原理的好办法，但实际上，它们都经过数学的界定。根据一个位置的经纬度，通过一套公式即可计算出投影坐标系上的一个对应点。绘图人员在建立地图投影时有多种不同的选择，它们各有各的特点。如何将可展表面围绕到地球模型上，如何安排方位，如何设计想象光源的位置只不过是众多参数中的几项而已。

A2. 12. 遗憾的是，没有任何一种方法可以完美地把球体坐标表示在平面地图上。也就是说，没有任何一种地图投影法能够全面地满足各种目的。每一种方法仅在保持某些特性方面较好，而其他方面则十分糟糕。投影方式的不同会引起效果不同的失真。因此，地图投影法可以根据各自的特征进行分类，其中主要有：

- **正确的面积。**大多数的投影都会在地图上造成拉伸效果。而且这种拉伸在地图上的各个部位不是均匀的，例如，靠近世界地图极点处的地形一般看起来较大，大于赤道附近的地形。例如，阿拉伯半岛要比格陵兰岛大出几十万平方公里，但在很多地图上，格陵兰岛看起来比阿拉伯半岛大好几倍。地图上正确表达所有各种地块相对面积的方法称为等面积投影法。莫尔威德投影法就是一个实例。
- **相等的距离。**没有任何一种投影法可以把地图上两点之间的距离表达得准确无误。记住这一点很重要，因为地理信息系统数据库的一个常用功能就是计算距离。对于一个小地理区域的大比例尺地图来说，造成的误差通常可以忽略不计。然而，对于一个国家或一个大陆的小比例尺地图，地理信息系统计算出的距离则不太可靠，除非系统可以根据比例尺进行欧几里德距离计算，对误差进行补偿。即使等距投影也不能保证所有距离都能正确显示，只是能够正确表达而已，即正确表达某一点或两个点到所有其他各点的距离，或沿着一条或多条线的距离。等距圆锥投影法即是一例。但必须注意的是，要想非常准确地计算距离，通常只能采用精确的几何式，而不是简单的欧几里德距离。这种计算基于经纬度坐标，即所谓的“大圆距离”。

- **正确的角度。**保形投影可以保证围绕所有点的角度以及小面积的形状不变。子午线与纬线以正确的角度相交。这些投影法对于航海或航空十分有用。墨卡托投影法即是一例。

A2.13. 因此，所有地图投影法都是在各种所需制图方法中寻求妥协。对于任何具体应用来说，总能找到比另外一种方法更为适合一些的投影法。除了投影法本身的特点之外，还要考虑到制图区域的大小、原始范围(比如南北方向与东西方向)的面积，以及该地区在全球中的位置(例如靠近极地、中纬地区还是赤道)。

A2.14. 地图教科书和很多地理信息系统手册都有全面详细的列表，说明哪种应用选用哪种投影方法最好。有时，最好的选择并不意味着能完美地保持任何特征。例如经常用于世界地图的罗宾逊投影法是一种折中的解决方案，其设计意图基本上出于美学考虑，比如为了制作地图册。在其他情况下，由于只绘制相对较小区域的地图，因此，任何一种投影法造成的失真对具体应用而言，均可忽略不计。

A2.15. 图A2.6说明了一些广泛使用的地图投影法。在图的上部，地球显示为球形，还有一种非投影经纬度坐标法，地图似乎是在平面坐标上绘制成的。很多地理信息系统数据发布软件发布的数字化地图数据有时是这种非投影式的“地理”坐标，用户通常能够直接把经纬度坐标转换成任何一种地图投影系统，但有时候，从某种地图投影转换成另一种投影十分困难。

#### D. 更为精确的制图方法：地理基准法

A2.16. 在从球形经纬度转换成平面坐标的过程中，由于地球本身不是一个具有同一半径的完美球形，因此这种转换更为复杂。精确的测量表明，地球表面处于高度变化和持续改变之中。更重要的是，地球是扁平的，地球中心到北极(短半轴)的距离小于地球中心到赤道(长半轴)的距离。为了更为精确地绘制地图，要把地球更精确地界定为一个椭圆球或扁球，极点半径与赤道半径符合一种特定关系(见图A2.7)。在地图使用的基准坐标中，用于说明椭圆球和原点及方位的参数称为“大地基准”(根据地球测量学的测地学命名)。

A2.17. 与椭圆球最近似的参数在地球的不同之处变化极大，因此界定出数百套不同的基准。不过，每个国家的地图制图机构在所有制图和地理空间活动中只采用一种标准基准，而绘制地区、大洲或世界地图时也仅适用少数几种标准基准。如果地图测绘部门更改了标准基准，会造成很多麻烦。近两个世纪以来，基准一直在不断重新界定，老地图基于过去的标准，而新的地图则可能使用新的和更为准确的标准。

A2.18. 对于覆盖大片地区的小比例尺地图或为某一具体目的绘制的草图来说，无需很高的准确度，不同基准造成的问题尽可以忽略不计。但对于大比例尺的精确地图，其偏差就非常大了。表A2.1列出的是根据横向墨卡托(UTM)通用坐标系得出的纽约联合国总部大厦的坐标，对于墨卡托坐标系下面还要详

图 A2.6  
常用的地图投影法

全球(从太空看地球)



非投影(地球或纬度/经度)



等面积圆柱投影法



墨卡托投影法



阿尔伯斯等面积圆锥投影法



等距圆锥投影法



兰伯特等面积方位投影法

赤道位置



倾斜位置



极顶位置



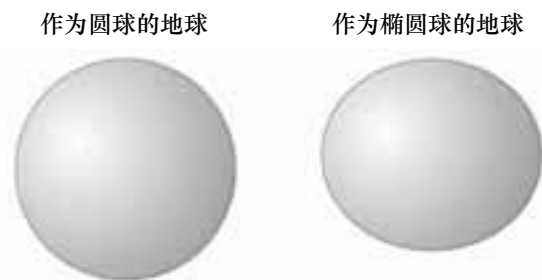
莫尔威德等面积投影法



罗宾逊投影法



图A2.7  
圆球与椭圆球的对比



细讨论。联合国总部大厦的经纬度坐标是利用不同大地基准投影到同一投影图中的。直到最近，较老式的克拉克椭圆球在美国一直作为标准适用。使用克拉克椭圆球与使用世界大地测量系统(WGS)所得出的坐标值，南北方向的地面偏差大约为300米，或者说在一幅1:25 000的地图上相差1厘米。把地球按绝对球形计算和按椭圆球计算的偏差值超过18公里。

表A2.1  
纽约联合国总部大厦按不同基准椭圆球投影后得出的坐标

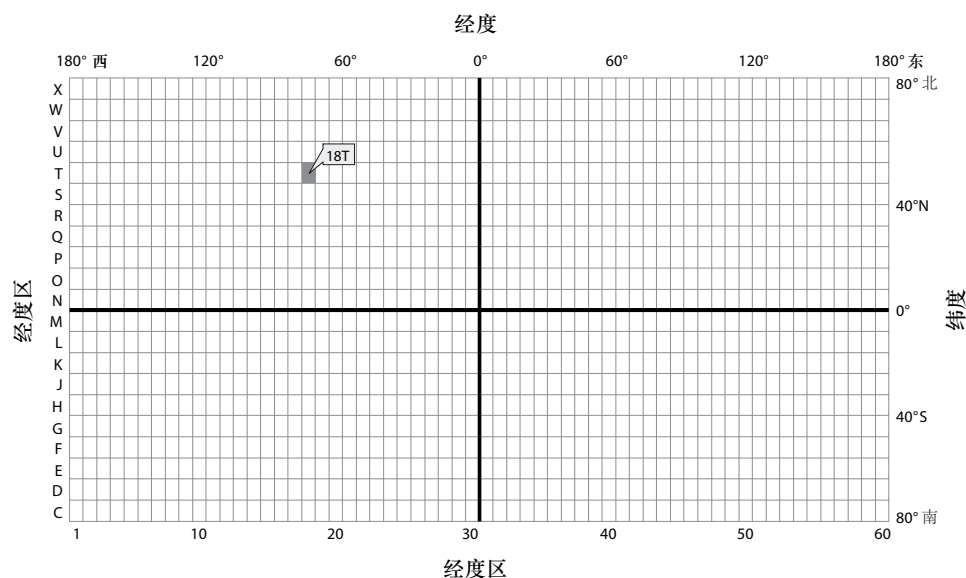
基准椭圆球	横向墨卡托通用坐标(米)	
	东向(x)	北向(y)
克拉克, 1866	587 141.3	4 511 337.1
克拉克, 1880	587 142.6	4 511 245.1
WGS84	587 139.0	4 511 549.7
贝赛尔	587 128.5	4 511 095.4
圆球	586 917.2	4 529 920.6

### 通用横向墨卡托基准体系

A2. 19. 通用横向墨卡托基准是一个值得详细讨论的地图基准体系，是绘制大比例尺世界地图时最常使用的体系之一。它基于横向圆柱投影(横向墨卡托)，其中柱面与地球模型沿子午线接触。对于世界不同地区，可以选用不同的“本地”子午线，因此在比例、形状以及正切线上的距离失真均很小。全世界的横向墨卡托通用基准体系包括60个经度区(见图A2.8)。

A2. 20. 每一经度区的宽度为六度的经度，正切子午线两边各为三度。横向莫卡托通用基准区从西向东顺序编号，其中1号经度区的范围为180°W到174°W，中心子午线为177°W。该经度区进一步划分为行，每行的高度为8度。这些行从南到北以字母编号，最南端为80度，编号为C，依次类推。由于两极的失真非常严重，因此在上述范围之外，没有定义通用横向墨卡托基准经度区。

图A2.8  
通用横向墨卡托基准体系



A2. 21. 坐标的计算以米(或英尺)为单位, 从中心子午线测量。沿东西方向的, 取东向值; 沿南北方向的, 取北向值。为了避免出现负数, 在东向值上加上500 000。出于同样道理, 在北向值上加上10 000 000, 但是这项添加值仅在计算南半球的坐标时使用。这种偏移量, 称为“伪东向”和“伪北向”。

A2. 22. 为了说明UTM系统的用途, 图A2.9给出了一个实例。纽约联合国总部大厦位于北纬 $40^{\circ}45'01''$ 和西经 $73^{\circ}58'04''$ 。该位置处于UTM的18T经度区内, 范围是西经 $72^{\circ}$ 至 $78^{\circ}$ , 北纬 $40^{\circ}$ 至 $48^{\circ}$ 。其UTM的x和y坐标以米计算, 分别为587 139.0和4511 549.7。这就是说, 联合国总部大厦大约位于UTM经度区的第18号( $75^{\circ}W$ )中心子午线以东的87公里, 赤道以北的4 512公里处。

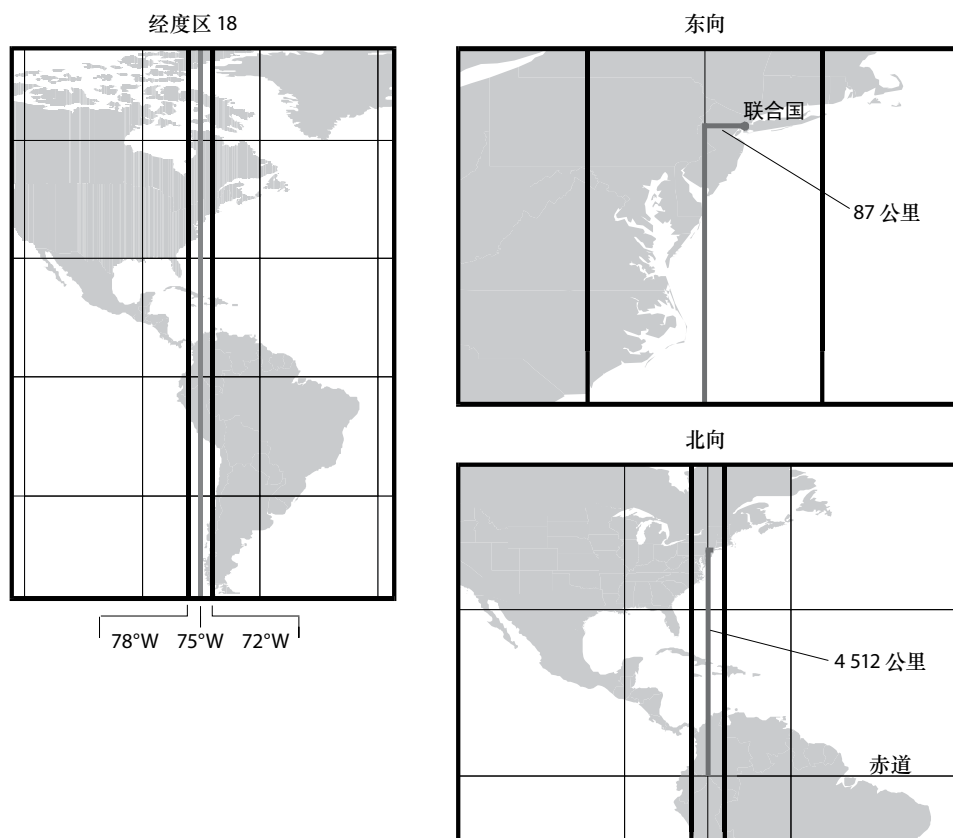
## E. 地图比例尺

A2. 23. 出版印刷的地图在覆盖面积方面的差别极大。国家或地区一级的地图上只能显示一些最重要的地理特征, 地方性地图则可显示更多的细节, 例如具体的房屋或溪流等。标准地图的覆盖面积按选定的制图比例尺确定。这个比例尺以地图上的距离与真实世界的距离之比的形式出现, 例如, 在一幅比例尺为1:25 000的地形图上, 地图上的1厘米代表真实世界25 000厘米或250米的距离。

A2. 24. 鉴于地图的比例尺是一个分数或比率, 它所代表的地面距离越大, 该地图的比例尺则越小。例如, 一副比例尺为1:1 000 000的地图就是一幅小比例尺的地图, 因为1除以一百万后得出的是一个非常小的数值(0.000001)。相比之下, 一幅比例尺为1:5 000的地图则是一个大比例尺的地图, 因为1除以5 000后得出的数值(0.0002)相对较大。因此, 小比例尺地图显示的是很大的面积, 而大比例尺地图则强调的是一个很小的区域。实际上, “小比例尺”和

图A2.9

联合国总部大厦在UTM基准体系中的位置



“大比例尺”经常混淆，因为按口头上的说法，“大”和“小”指的是覆盖面积或一个物体的尺寸，而不是分数本身的大小。例如，全球气象模型经常叫做大比例模型。所以，为了避免出现误会，可以将其直接称为“制图比例”。

A2. 25. 某些常用的地图比例尺如下：

地图比例尺	地图上1厘米所代表的实际距离	
1:5 000	50 米	较大比例
1:25 000	250 米	
1:50 000	500 米	
1:100 000	1 公里	
1:500 000	5 公里	
1:1 000 000	10 公里	较小比例

A2. 26. 假设从模拟地图向数字化地理数据库转换，非常重要的是要强调，实际上，数字化的地理数据是没有比例尺的。地理特征坐标一旦输入到地理信息系统中，它们就可以按任何一种指定的比例显示。用户在处理数据的过程中，可以把地图放大和缩小，不露丝毫痕迹地迅速转换比例。尽管如此，重要的是需要记住，这些数据通常有其物质来源(纸质地图、图像等)，而这些原

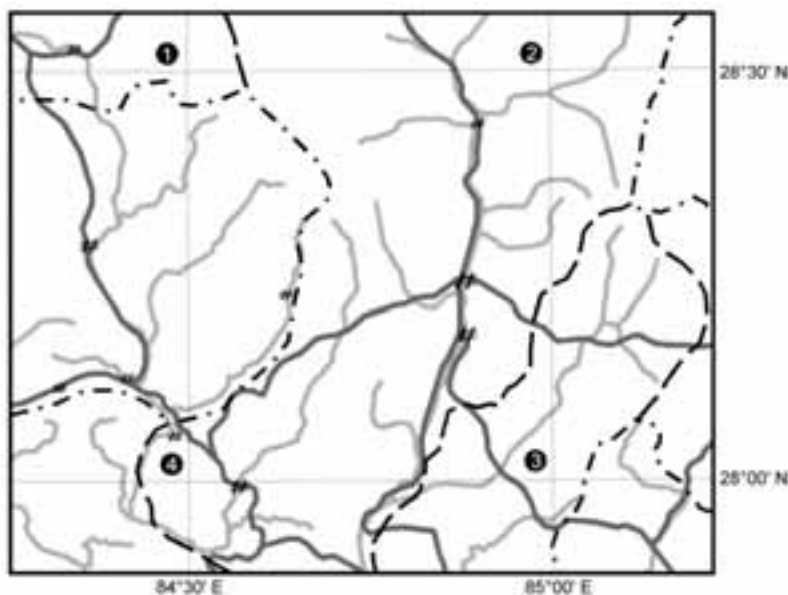
始材料本身都是有比例的。例如印刷式地图，其比例尺不同，显示的详细程度也有很大的不同。使用1:25 000比例尺的地图可以显示出构成村庄的每座房屋，而在一幅比例尺为1:500 000的地图上，整个村庄，如果真能显示出来的话，只是地图上的一个点。

A2.27. 把地图上的地理特征加以简化或加总的过程称为“泛化处理”，这是地图绘制中的一个重要组成部分。地理特征经过泛化处理之后，蜿蜒曲折的乡间小路变成了直线，区域边界上的细节消失了。由于泛化处理的结果，对比例尺为1:250 000的地图进行数字化处理，然后打印成一幅比例尺为1:5 000的地图就变得毫无意义了，同样，把取自比例尺差别较大的地图的数据集进行合并的努力也是徒劳的。这说明，在一个数字化地理数据集的文件记录中，标明地图的原始比例是极其重要的。此外，由于存在上述种种与地图比例尺有关的问题，在开展一个大型的数字化地图绘制项目时，应尽早确定成品输出的比例要求，这样才能使数据库的开发工作建立在足够的资料来源的基础上。

## F. 地理基准实例

A2.28. 一幅地图经过数字化处理或扫描之后成为合适的地图单位坐标存储到地理信息系统内，与其有关的地理基准问题已在本《手册》第二章讨论数字化地图集成过程时论述过。为了说明地理基准问题，下面将通过一个具体实例进行说明。图A2.10所示的地图经过数字化后形成若干个数据层。数字化之后，坐标以数字化的表格单位为基准，此处为英寸。对于该地理区域，为了与其他数字化数据一起使用数字化地图，需要将数码化表使用的坐标转换成真实世界坐标，即对应于该地图的原始投影。不熟悉坐标系和地图投影的读者，可参阅上文A2.1-A2.7段。

图A2.10  
地图上的控制点





A2. 29. 第一步是确定合适的控制点，这一步骤通常是数字化处理过程的一部分。控制点应恰当地分布到所需的区域上，以改善参数转换的估算值。这就是说，它们不应放在一个区域内，更不应仅仅集中在地图的中央。除了道路、河流、行政机关以及城镇外，地图上要显示出一个由经纬线构成的、间隔为半度的规则网格。这个所谓的经纬网的交叉点正好作为合适的控制点，因为它们的坐标值很容易确定。在这张地图上，将四个选中的控制点按1至4顺序编号。它们的坐标值分别为(84.5, 28.5)、(85.0, 28.5)、(85.0, 28.0)和(84.5, 28.0)。注意，由于地理信息系统程序使用的是平面坐标，所以必须把每对坐标值规定为经度/纬度(即x/y)而不是纬度/经度。出于同样理由，需要把坐标规定为十进制的度数，而不是纸质地图或地名字典中通常采用的度、分、秒格式。

A2. 30. 遗憾的是，我们不能直接使用经度/纬度坐标来进行变换，因为纸质地图原本不是按照经度/纬度平面标注的，只有非常少的地图采用这种标注方法。事实上，纬度和经度网格并不包含直线，根据这一事实即可断定它们不属于平面坐标系。在本例中，地图的原始投影是阿尔伯斯圆锥等面积投影，其参数如下：

- 标准平行线：北纬27°和30°
- 中心子午线：84°
- 原点纬度：28°

A2. 31. 这些地图参数通常在地图的页面上标明。在进行坐标变换之前，首先需要把控制点的经/纬坐标转换成阿尔伯斯投影时的真实世界坐标。在大多数软件程序中，转换方法是列出每一对的经度/纬度数(因为经度是x坐标，而纬度是y坐标)，然后放入文本文件中完成坐标，或者，通过菜单界面，在系统投影转换模式中规定适当的投影参数。

A2. 32. 当然，如果能够直接从地图上读取真实世界的控制点坐标，这项额外步骤就没有必要了。例如，地形图如果以UTM投影为基准绘制而成，即有这样的可能。还有一种情况是控制点已经通过全球定位系统加以确定，因为全球定位系统可以自动把坐标转换到特定的地理投影之中。

A2. 33. 现在，数码化表上有四个控制点的坐标对，以及真实世界的投影坐标，在本例中，单位为米。两组坐标值都列在表A2.2中。例如，第一个控制点位于中心子午线(84°E)以东49公里和原点纬度(28°N)以北55.5公里处。

A2. 34. 第三步是根据两组坐标值进行参数转换计算，大多数的地理信息系统软件都有这项功能。从技术上说，这些参数是通过以下回归方程估算出来的：

$$x' = a + bx + cy$$

$$y' = d + ex + fy$$

其中： $x'$ 和 $y'$ 是真实世界坐标，而 $x$ 和 $y$ 是数字化仪上的控制点的坐标； $a$ 、 $b$ 、 $c$ 、 $d$ 、 $e$ 和 $f$ 是要估算的参数。在转换过程中，估算误差就是回归计算得出的余数。

A2. 35. 表A2.2列出了每一个控制点在输入坐标系(数字化仪使用的单位)和输出坐标系(以米为单位的阿尔伯斯投影)中的坐标对。另外,表中还列出了转换误差(余数),误差是经过计算得出的,输出坐标的单位为米。可以看出,转换误差在x方向约7.8米,在y方向约14.6米。这些误差几乎从来没有为零。误差来源包括地图纸张收缩变形或折叠而造成的失真,以及对控制点坐标进行数字化处理时出现的测量误差。如果一个或多个控制点出现了非常大的误差,一般说明出现了重大错误,例如把x和y坐标弄颠倒了,或者控制点的标识符出现了错位。总之,这一步骤要非常仔细小心,因为它会极大地影响最终结果的准确度,从而影响最后的地理信息系统数据库的可用程度。

A2. 36. 该表还提供了一项指标,即转换的总误差,这是一项均方根误差,以输入和输出坐标系单位表示(分别为英寸和米)。出现很大的均方根误差表明,以输入和输出地图上的单位表示的控制点的位置与相对位置不对应。对于一个大型数据转换项目来说,应该规定一个可接受的最大均方根误差范围,并维持在此范围内,何种程度可以接受,这取决于原始纸张地图的比例以及应用方面的准确度要求。人口普查地图不需要很高的准确度,而地籍方面的应用必须满足相当高的标准。

表A2.2  
参数转换

控制点	坐标值,按数字化仪的 单位(英寸)		坐标值,按投影的真实 世界的单位(米)		坐标值误差,按真实 世界的单位(米)	
	x	y	x	y	x	y
1	11.777	19.660	48 936.2	55 529.6	-14.59	7.80
2	26.670	20.661	97 871.5	55 835.2	14.60	-7.81
3	27.696	3.824	98 333.0	409.3	-14.55	7.78
4	12.751	2.810	49 166.9	102.3	14.54	-7.77
均方根误差(输入、输出)			0.005034 16.524			

A2. 37. 系统将在同一步骤中,把地图数据库内的所有坐标转换成输出坐标系。然后,输出数据库将以原始纸张地图的坐标作为正确基准。此后,可将地图投影到不同的投影中,例如,根据一种不同标准投影,将其集成到综合数据库中。这里只是简要地说明了坐标转换的一般原理。尽管在实际工作中,这一切都是由软件实现的,但对地理基准所涉及的各个步骤有所了解,有助于理解该步骤的重要性。

## G. 实际考虑

A2. 38. 任何一项大型数字化地理数据库工程(例如支助人口普查地图)都需要把很多不同来源的地图信息集成在一起。因此,需要选择一个标准的投影和坐标系。理想的情况是,选择的基准体系应该与国家其他制图活动使用的基准体系相同。大多数国家使用的标准投影和坐标系都是根据国家的领土条件,针对不同比例尺的地图精心选择出来的。

A2. 39. 几乎所有地理信息系统软件都有坐标转换功能，把一个基准体系的坐标转换成另一个坐标系(例如，从米转换到英尺，或从数字化单位转换到地图单位)，从纬度/经度转换到一种地图投影，或在不同的投影之间进行转换。这些软件还允许用户选择一种大地测量基准以及其他关联参数。在有些罕见的情况下，某种特殊投影得不到普通软件的支持，这时则需使用特殊投影软件。本《手册》第四章详细讨论过的全球定位系统也支持所选的地图投影以及最常用的大地测量基准。因此，实地工作期间，既可按纬度和经度对，也可按某种投影体系收集坐标。

A2. 40. 投影和基准信息通常包含在地形图内，然而，对于数字化的地图数据集来说，有一个问题是，标准的地理信息系统格式并不需要明确存储投影信息。例如，人口普查机构可以获得关于道路或水文方面的地理数据集，但却得不到地图投影方面的信息。如果把这样的数据与数字化人口普查地图合并在一起，它们之间的匹配不会十分完美。因此，不可能进行纵向集成，除非将两种数据集都纳入相同的投影系统。如果无法根据数据集谱系确定地图的原始出处，惟一的选择就是临时协调两套数字化地图，但结果可能出现严重误差。由此可见，在保存数字化地图数据集时，正确保存好数据集文档和元数据(与数据有关的信息)是何等的重要。

A2. 41. 最后一个需要实际考虑的问题是不同经纬度存储格式之间的相互转换问题。经纬度通常按度、分、秒的格式表示。例如，纽约联合国总部大厦的位置是北纬40°45'01"，西经73°58'04"。为了把这个经纬度坐标值输入到地理信息系统或地图投影系统中，首先需要把坐标值转换成十进制的度数，也就是说，让它们看上去与普通的x和y式的笛卡尔坐标相同。以下以联合国总部的经纬度为例，说明如何把度、分、秒转换成十进制度数。

$$40 + \frac{\left(45 + \frac{1}{60}\right)}{60} = 40.7502778$$

$$73 + \frac{\left(58 + \frac{4}{60}\right)}{60} = 73.9677778$$

A2. 42. 由于联合国总部的经度位于格林尼治子午线以西，因此在十进制度数中以负数(即，-73.97)形式出现。同理，位于南半球的纬度数也要以负数表示。

A2. 43. 再将纬度转换成度、分、秒，过程如下：

度：40

分：0.7502778 × 60 = 45.016668 = 45

秒：0.016668 × 60 = 1

## 附件三

### 数据模型操作

#### A. 引言

A3.1. 本附件讨论地理数据模型操作问题，包括一个数据字典详细内容的实例。人口普查部门通过数据字典可为用于人口普查的地理数据库制作文档。一个简单的、与地理人口普查产品一起发布的数据字典见附件四。

#### B. 关键术语的定义

A3.2. “空间数据模型”系指对地理实体以及地理实体之间关系进行的描述。所谓地理实体指房屋、行政单位以及河流等。在面向对象的数据模型中，其定义通常还包含对实体进行的操作。数据模型独立于任何具体软件包，因此，用户可以在任何综合性地理信息系统软件包中进行数据模型操作。

A3.3. “空间数据结构”系指建立一个具体的数据模型，包括用于表示不同实体类型的具体文件结构。例如，行政单位或水域用多边形表示，而多边形就是一系列坐标点，其中第一个和最后一个是一个坐标点。有了数据结构，界定地理实体关系的软件才能运行。例如，一条道路恰好与一个多边形边界的一部分重合，而该多边形定义的是一个行政单位。

A3.4. “数据格式”是一个使用较为普遍的术语，通常指软件系统中具体一套数据结构。有些商业数据格式，由于得到广泛使用，因为成为一种事实上的标准。例如DXF格式(绘图交换格式)最初是为了开发AutoCad软件包制定的，但现在几乎得到了所有商业地理信息系统软件包的支持。

A3.5. “数据字典”是一个详细说明数据模型以及所有实体及其属性标定代码的主文档。

A3.6. 最后，“数据库模式”对空间实体、属性表以及界定完整综合空间数据库的集成规则进行描述。

#### C. 模板实例

A3.7. 下列模板实例取自《加拿大国家地形数据库》的《数据字典》，其中包含了非常全面的地理实体定义(加拿大测绘学，1994年)。

表A3.1  
用于定义空间数据模型的汇编信息

实体名称	地理特征的简称。
定义	地理实体的详细描述。
定域属性	此类属性仅有有限几个预定义值，例如行政单位类型(区、省、等等)或者路面类型等。这些预先确定的代码是可能值的“域”。
变域属性	<p>此类属性具有无限多的可能数值，因此它们的域不能限定，例如行政单位独有的标识符、单位的人口数量或河流名称等。</p> <p>每一属性由以下信息加以说明：</p> <p>名称</p> <p>类型，例如字母数字型(A)、整数型(I)或实数型(R)</p> <p>允许使用的字符或数字个数</p> <p>值域——定域属性的所有可能值及其定义列表；对于变域属性而言，即属性定义。</p>
授权属性值的组合	<p>对于定域属性，列出所有允许的属性组合。例如，对于行政单位来说，只有地区和省可能有正式的行政首府。所以，如果行政单位不是地区或省，列有首府名称的另一属性应保持空白。关于授权属性值组合的信息，对一致性的自动检查过程十分有用。</p> <p>如果该实体没有定域属性，则填入“无”。如果只有一个定域属性，所有的授权值都要列出。如果有几个定域属性，则列出所有的授权组合值。</p>
关系	<p>系指对地理实体与其他空间特征关系的描述。这一功能十分有用，例如，用来界定河流或道路如何与行政单位边界或点查区边界相重合。关系由以下几种特点界定：</p> <ul style="list-style-type: none"> <li>— 实体名称和实体几何形，如点(P)、线(L)或面积(A)。</li> <li>— 关系，例如，连接，一条线连接到一个点；或者，共享，一个区域与一条线共享一条边界。</li> <li>— 基数，由一对数值定义，确定一个实体可以涉及关系的最少和最多次数。例如，一条道路的交叉点与道路特征发生关系。在交叉点处至少有一条道路与其相连，从理论上说，最多可以有无限条道路与其相交。如果最大数不能确定，则以N来表示。因此道路交叉与道路的关系就是(1, N)。</li> <li>— 相关实体的名称和几何形。</li> </ul> <p>注意：这里仅仅指地理特征之间的关系。在地理属性表和外部表字段之间的关系需要另外分别界定。</p>
几何表示与最小尺寸(米)	用于表示实体的几何特征。对于行政单位，几乎总是多边形。然而，对其他地理特征而言，用何种几何元素表示空间实体取决于地图的比例尺。例如，一个村庄在大比例尺的地图上(如1:25 000)的面积可表示出它的轮廓，但在小比例尺的地图上(如1:250 000)可能成为一个点。在同样的比例尺下，大的村镇可能以面积表示，小的村落则可能以点表示。随着地理特征类型的不同，实体的最小尺寸可能指它们的“表面积”、“宽度”、“长度”或“高度”。
说明	定义实体所需的各项附加信息以及其他描述字段内的脚注。
图解	为了说明一个实体模型化的方式，用一个图表明该实体与其他各种实体之间的关系。

A3.8. 数据库模板中最为重要的信息是对每一实体所做的定义，以及对所存储的地理特征所有属性的详细说明。对于很多人口普查地图绘制项目来说，基本的数据库描述符可能就足够了。然而，尤其是对于一个将被纳入国家

地理数据库的人口普查数据库来说，建议在数据库设计时应多付出一些时间和精力，确保能够与其他机构的信息兼容。在这种情况下，行政单位或人口普查部门与其他地理特征之间的关系应当明确加以界定。

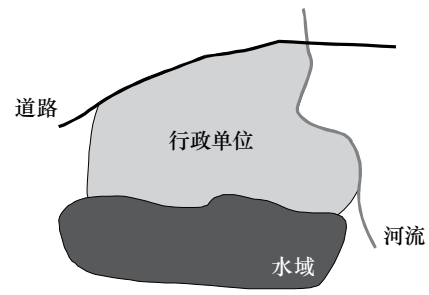
A3.9. 为了说明数据字典的内容，表A3.2给出了一个实例，说明了如何对行政单位数据层进行定义。这个例子仅用于说明目的。具体的规定将随各国的具体情况而有所不同。应该指出的是“定域属性”是离散的，而不是连续的。

表A3.2

实例：具有三级管理的国家行政单位

行政单位	
具有法定边界的地理区域，设置的目的在于实行政管理或其他政府职能。	
定域属性	
行政单位类型I(1):	
1—省份	第一级行政单位
2—地区	第二级行政单位
3—地方	第三级行政单位
城/乡指标I(1):	
1—不适用	仅在地方一级区分农村和城市
2—农村	一个具有支配性农村特点的行政单位
3—城市	一个包含镇或市的行政单位
变域属性	
行政单位标识符I(14)	
说明：在本数据库实例中，所有的属性信息(例如名称、别名、住户数量、人口等)均分别存入数据表中，这些数据表通过行政单位标识符链接地理属性表。	
授权属性值的组合	
省份(不适用)	
地区(不适用)	
地方(城市)	说明：只有这些组合是可能的。例如，没有城市省份或农村地区。
地方(农村)	
关系	
行政单位(P)	共享 (0,N) 道路(L)
行政单位(P)	共享 (0,N) 河流(L)
行政单位(P)	共享 (0,N) 水域(P)
说明：道路与河流以线(L)表示，道路、河流可能与行政单位的边界重合，行政单位以多边形(P)表示。同样，行政单位的边界也可能与水域共用一条边界，例如一个用多边形表示的湖。(0,N)指的是关系的基数。其含义是最少可能有零条道路与一个行政单位的边界重合，而最大值不确定(N代表任何数)。	
几何表示与最小尺寸	
如果一个行政单位的面积小于或等于1平方公里，以一个点特征表示；如果大于1平方公里则用一个面积特征表示。	
说明：行政单位必须与点查区边界重合。行政单位应覆盖整个国家的全部领土。换言之，不能有任何一块国家领土没有分配到行政单位中。	

图 解



## 附件 四

### 用于发布的数据字典实例

A4.1. 下面是一个数据字典的实例，由假想的国家波普兰迪亚发布的地区性人口普查地理数据库。数据字典这个词有时可与元数据这个词交换使用，尽管在实践中，这两个词并不相同。数据字典的日期早于元数据，通常“数据字典”一词指的是与机构发布的数据一起发布的信息。国际标准化组织的元数据标准提供了具体的指导准则。各国可以使用“界定”方式来改变通用的元数据格式，供本国使用。下面的实例仅用于说明目的。真正的数据字典的内容应当由国家人口普查部门精心设计，充分考虑到与国家有关的各种问题。

#### 数据字典：地方性人口普查地理信息系统数据库

数据库名称	波普兰迪亚地方性数字化地理人口普查数据库
数据来源	国家统计局(NSO)、人口普查局、地图部(1996年)，1995年波普兰迪亚国家人口与住房普查。
数据库内容	<p>该数据库包括一个用于全国的地方性地理数据层。地理信息系统数据库的文件发布格式为Arc图形文件格式(环境系统研究所公司)、MapInfo交换格式(MapInfo, Inc.)，或普通坐标值文本文件。本文档参照Arc图形文件版本。</p> <p>地方性地理信息系统层的地理属性数据表(LOC.DBF)仅包括基本信息，如地区编码(LOC_CODE)以及属于其中的行政单位的名称。两个外部数据表与地理信息系统数据库一起发布，一个包含人口普查得出的人口特征值(POP.DBF)，另一个为住户属性(HH.DBF)。通过诸如LOC_CODE的公用字段，可将两个数据表链接到地方性地理信息系统数据库。</p> <p>除另有说明外，所有的数据均以2005年7月1日的人口普查结果为准。</p>
行政管理与报告单位	数据库包括9个省123个地区的1 291个地方的信息。
软件和硬件要求	可通过能够导入Arc图形文件或MapInfo交换格式文件的任何地理信息系统或桌面地图绘制软件查看数据库。系统的最低配置取决于数据访问软件。一般来说，486Mhz或更快的IBM兼容的个人计算机以及至少8mb的RAM内存即可满足需要。该数据库可通过只读光盘驱动器或计算机上的硬盘驱动器进行访问。需要16mb的硬盘空间。
数据库发布格式	数据库以非压缩形式通过只读光盘发布，可直接对数据库进行访问。
投影法	等距离圆锥投影法
标准平行线	北纬20°和北纬60°
中心子午线	西经140°
坐标单位	米
坐标偏置	无



数据库名称	波普兰迪亚地方性数字化地理人口普查数据库
源地图比例尺	各种不同的比例尺。大多数城市地方的地图使用1:25 000或更大的比例尺；农村地区的地图使用1:50 000或更小的比例尺。
一般准确度信息	根据国家制图机构的信息，坐标平均估计准确度为农村地区+/-100米，城市地区 +/-30米。
彼此不连接的报告单位	某些地方包括不止一个多边形。属性表包括一个字段(FLAG)，对于主要多边形，数值为1(只包含一个多边形的地方)，对于次要多边形，数值为0。为了避免人口普查数据加总时出现双重计数的情况，需要先把FLAG值为1的地方选出之后再行加总操作。
相关产品	国家统计局办公室曾公布过类似的用于点查区的数字化地理信息系统数据库。因为点查区的数目非常多，有为每个省分别建立的地理信息系统数据库。欲知更多情况，请联络国家统计局办公室。
参考资料	国家统计局办公室(2005年)，《2005年波普兰迪亚国家人口与住房普查，关于人口普查制图活动的技术报告》，人口普查局，地图部。 国家统计局办公室(2005年)，《2005年波普兰迪亚国家人口与住房普查，方法与管理报告》，人口普查局。 国家统计局办公室(2005年)，《2005年波普兰迪亚国家人口与住房普查结果》，人口普查局，地图部。
联系信息	国家统计局办公室，人口普查局 地图部，用户服务部 邮箱：9999 塔罗塔，桑姆巴斯省 电话：99-99-99999 传真：99-99-99998 电子邮件：geog@census.gov.xx 网址：www.census.gov.xx

## 地理数据文件

### LOC.SHP——地方边界地理信息系统数据库

文件名:	C.SX		
文件类型:	ESRI A.图形文件		
特征类型:	多边形		
有关文件:	C.BF	多边形属性表(图形文件的一部分)	
	-BF	普查人口指标	
	-BF	普查住户指标	
	C.SX	A.使用的内部地理索引文件	

## 属性数据文件

### LOC.DBF: 地方性特征值

字段名	描述	字段规定	范围	编码	丢失值
C-E	正式地方编码。用于链接外部数据表pop.dbf和hh.dbf。该地理编码的构成方式是连接行政单位的标识符：省2位+地区3位+地方3位。	整数，8	正值	无	-999
AREA	以平方公里计的地方的面积。	实数，6.1	正值	无	-999

字段名	描 述	字段规定	范 围	编 码	丢失值
F3A	表明该多边形是否是该地方的主要多边形。对于含有两个或多个多边形的地方，其中只有最大的一个或最重要的一个的值为1。	整数，1	0-1	0-次要 1-主要	
RBAN	表明该地方是在城市还是农村。	整数，1	0-1	0-农村 1-城市	-1
CNA E	地方名	字符，25	无	无	“n.a.”
ISTNA E	地区名	字符，25	无	无	“n.a.”
R-VNA E	省份名	字符，25	无	无	“n.a.”
AREA-TA	地方总面积，单位：平方公里	实数，10.3	正值	无	-999
AREAN	地方陆地面积，单位：平方公里	实数，10.3	正值	无	-999
AREATER	地方水域面积，单位：平方公里	实数，10.3	正值	无	-999

POP.DBF——普查人口指标

字段名	描 述	字段规定	范 围	编 码	丢失值
C-C-E	正式地方编码。用于链接外部数据表	整数，8	正值	无	-999
T-T	点查人口总数	整数，7	正值	无	-999
ENS	人口密度，每平方公里人数(T-T/AREA)	实数，5.1	正值	无	-999

HH.DBF——普查住户指标

字段名	描 述	字段规定	范 围	编 码	丢失值
C-C-E	正式地方编码	整数，8	正值	无	-999
NN	住户数量	整数，7	正值	无	-999
EAE	户主性别	整数，1	0-1	0-男 1-女	-1



## 附件五

### 专题地图设计

#### A. 引言

A5.1. 本附件简要介绍制作专题地图时在设计方面的考虑。不过，本概述不可能包含围绕地图的信息内容的所有问题；如有必要，应该查看教科书。绘图人员把地图分成若干类型。普通地图的作用是为确定方位提供基准框架，主要显示地面上可以观察到的实际地理特征。这些特征可以是自然形成的，如河流、山脉、海岸线；也可以是人工形成的，如道路、居民点等。基准地图也可以显示一些地面上看不到的内容，最典型的例子是行政边界以及用来表示经纬度的网格。地形图也可归入普通地图或基准地图之列。它们在绘制人口点查区地图时十分有用，因为它们能够提供地理特征信息，点查人员可以根据这些信息找到分派给自己的工作区。

A5.2. 更为适合表达人口普查结果的地图是专题地图。它们能显示地面上不容易直接观察到的物理或文化现象的分布情况。专题地图可以基于定性信息或定量信息。定性信息专题地图的实例如根据母语或宗教对人口分布情况进行说明的地图。定量信息专题地图有时也叫做统计地图，它说明的是地图所绘制的地理特征的相对尺寸。例如选择代表国家各个城市的符号时，要根据每个城市的大小进行选择。还有一个例子是，报告区要根据人口密度在地图上以深浅不同的色调表示。大多数为人口普查目的制作的此类地图都具有这种性质。

#### B. 地图设计原则

A5.3. 尽管地图经常用于分析目的，但是它并不擅长表达准确的数据值。在地图上，数据值被转换成一种直观符号。绘图人员必须把数据值划分成几个档次，以便获得易于管理的一定数量的类别，不同的类别用不同的颜色或符号表示，这就是说，某些信息在地图显示的过程中丢失。地图的优势在于显示趋势、相对数量以及指标值的分布情况。如果要了解准确的数值，则更宜使用那些可以通过数据库查询的数据表格或数字化地图。

A5.4. 制作说明性地图是制作人员将思想或概念传达给读者的一种设计过程(Monmonier, 1993年)。这种过程类似于其他通过图表、图像或视觉手段进行定性或定量信息交流的形式。因此，一些指导图形设计的原则同样适用于地图绘制。

A5.5. 简单、明了是最重要的设计原则。很多地图最后变得杂乱无章，因为绘图人员试图在一个很小的空间里塞入过多的内容。有一个概念十分有用，即最大化“数据油墨比”的塔夫悌原则(1983年)。我们可以将其用于地图绘制，意思是要将大部分油墨用在表现地理数据而不是绘制多余信息上，不要有画蛇添足之笔。标题上的“……地图”或“图例”没有必要，许多方框、边界线有时指北箭头和比例尺甚至也不需要。当然，和大多数原则一样，这一原则也有例外。为了便于理解，有些地图元素，如图例、简单的标题以及资料来源等也需要使用。

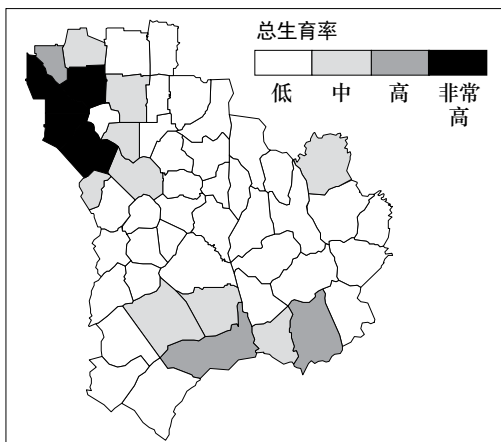
A5.6. 简洁还意味着不浪费空间。由于高分辨率的激光打印机已经十分普及，因此不必为显示所有细节而全部印制出大版面地图。地图设计得越好，打印的尺寸可以越小。为了节省空间，还要注意避免使用过大的字号、图例或内插图。

A5.7. 注意“视觉层次”是另一个重要概念。这个原则适用于地图元素以及所有组成部分的安排。对于地图本身，颜色和符号的选择能够反映出数据值的主次顺序。例如在一幅说明儿童死亡率的地图中，数值最高的报告区应使用最强烈的彩色或最重的灰度等级，这是应该立刻抓住看图者注意力的热点。例如在图A5.1中，从低到高的等级特意使用了由浅至深的灰度，突出“非常高”的一级。通过深色与周围浅色的对比，造成视觉上的层次。用深色调包围一个相对较浅淡的区域，也能产生同样的效果。下面我们来更详细地讨论一下选择地图颜色的问题。

A5.8. 绘图人员还可运用其他技巧把观察者的注意力吸引到地图上的特定区域。例如，把地图上最重要的特征边界勾勒清楚，使它们从背景中凸现出来。有时可利用注释或箭头指向特定部分，但这种做法常常会使地图显得杂乱。

图A5.1

通过选择颜色或灰度级构成视觉上的层次



A5.9. 对于整个地图的构成, 这些原则也同样适用。地图中最重要的部分就是地图信息本身、标题和说明符号的图例。这些应该是地图上最突出的特征, 而其他地图元素要慎重添加。

A5.10. 统计专题地图的色彩问题是一个多方面的问题(另见上文A5.4段)。绘图人员应该明白利用色彩进行直观展示与墨水及其他用品的额外成本之间的平衡。一个要考虑的问题是色盲。地图应避免使用红色和绿色来显示数据值, 因为色盲群体常常不能分辨这两种颜色。更笼统地说, 一幅彩色地图应该可以在黑白打印机上复制, 而信息内容不会丢失。这意味着要酌情考虑明暗浓度和取代图案填充。最后一项需要注意的事情与地图绘制者的责任有关, 即不要由于地图设计而触犯人口中的任何群体。绘图人员需要了解和掌握不同地区或人口群体的各种敏感特征。某些符号或颜色对国家的不同人种或种族群体会有一定的负面或正面含义。地图设计应避免使用与某些人口群体固定概念有关的各种符号。最重要的是, 绘图人员应该明白明智地而不是为没有必要的效果使用色彩的总体需要。

## 1. 专题地图的要素

A5.11. 专题地图由若干部分组成。地图本身包括底图, 显示诸如国家疆界等边界, 以及其他基准特征, 如主要河流和城市等。底图为读者提供了一个定位框架, 便于将国家的一部分变量与另一部分进行比较。第二大要素是专题地图的铺陈, 用以表现某一变量的地理分布。

A5.12. 除了实际地图信息之外, 一份符合出版质量的地图还应包括一些附加因素, 特别是:

- **标题和副标题。**应该简明扼要。
- **数据来源、可信度和制作日期。**使读者对地图产生可靠感和信任感。必须包括版权信息。有些机构定期出版的地图还可加上参考号和版本号供内部使用。其他与地图理解有关的说明信息也应添加上去。对于大版面地图, 投影参数也需标明。
- **图例。**说明地图中的变量如何转换成地图符号, 例如, 哪种颜色代表哪一范围的人口密度。不要忘记在图例中标明所用的测量单位, 例如, “每平方公里人数”。
- **地图比例尺。**便于用户在地图上测量距离。对于系列专题地图, 如人口普查地图册, 所有地图均按同一比例制作, 因此不必在所有页上都注明比例信息。同样, 对于那些尺寸较小、读者非常熟悉而不会去测量的地图也不必标注比例尺。使用图解比例尺比简单地写上一个数字要好(如1:1 000 000)。即使地图在复制过程中经过缩放, 比例尺图解仍然可用。反之, 用来绘制原图的标称地图比例尺这时就不正确了。

- **指北箭头。**如果所有地图都是以北定向的，基准图上就不一定必须标明指北箭头。尤其是众所周知的地理区域，如全国地图，则更没有必要添加箭头了。如果为了更好地配合页面安排，对地图进行旋转处理，则一定要标明指北箭头。
- **地图边框和边界线。**用于区分地图的不同要素。是否使用这些图形要素在很大程度上是一个地图设计问题。过多的线条和边框会使地图看起来杂乱无章。只有各种地图要素容易混淆时，才使用额外的边界线。
- **地名和标志。**用于辨认地理特征或统计区。
- **格子线。**经纬线网格(平行线和子午线)用于在地图上确定方位。在小比例尺的地图上应当标出。
- **定位图。**用于表现在主图范围内的位置。例如，显示地区人口密度时，可通过一幅小地图说明该地区在国家或省份中的具体位置。
- **内插图。**与定位图类似，不过不是显示主图所覆盖的区域的位置，而是以较大的地图比例显示地图的某个小的局部。例如，省级地图可能会附上一个小的内插图，显示省会地区，或更详细的小地区的信息。
- **文字和注释。**背景信息或说明，应做到简单扼要。
- **其他图形元素。**如表示变量统计分布的直方图或地图绘制单位的标识符号。

A5.13. 图A5.2和A5.3是两个含有很多专题地图要素的实例。图A5.2是一幅假想国家波普兰迪亚的一级行政单位图。衬托地图的是用于地理基准的经纬网格，另外将首都、省会和主要河流作为基准补充了进来。所有地理特征都恰当地标出，不同类型的地理特征采用不同的字体。地图下面有图解比例尺、说明专题特征的图例以及地图来源。如果有统计办公室标识，则可以添加到每幅地图上。没有指北箭头，理由有二：首先，没有异常的方向定位，经线清晰，图的上部为北；第二个理由不那么明显：所采用的地图投影法使经线向北方汇聚，也就是说，在不同经线上北方的方向略有不同。

A5.14. 图A5.3中的专题地图表示的是波普兰迪亚的一个省的人口密度。这类地图可以，比如说，在人口普查出版物中补充按省显示人口特征的表格。这样的地图也可以包含图表或其他图示。这份地图设计得相当简练。标题点明地图的专题，副标题说明了地理区域。图例没有使用大小一样和不同颜色的方格，而是按人口密度等级以直方图的形式表示。这种表示方法不仅达到了传统上利用颜色浓淡代表数值的目的，而且还表现出了该地区的数值的频数分布。对于更为复杂的地图，即包含更多区域的地图，还可以把属于每一类地区的实际数目添加上去。为保持画面的清晰和简洁，本地图未标注数字。在图例和数据来源之下，有一个小小的定位图，表示卡塔尼亚省在国家中的位置。在定位图内，一般不必添加国家标志，因为读者都能辨认出自己国家的形状。

图A5.2  
行政单位和主要城市中心地图实例



## 2. 测量级别和图形变量

### (a) 空间维度

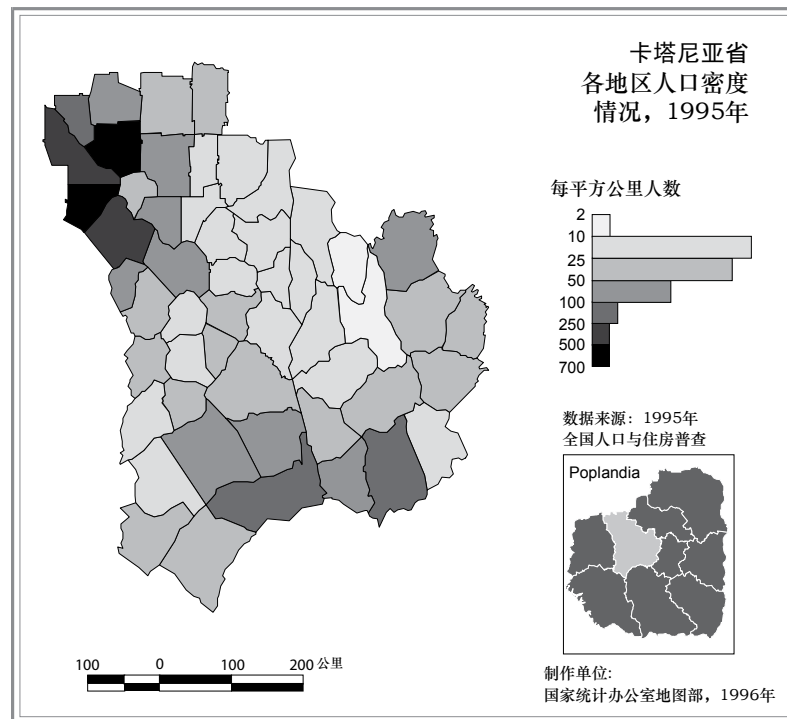
A5.15. 专题地图不仅显示地理特征的位置，而且提供有关地理特征的其他信息，即在每一个地理位置上的变量值。因此，专题地图不仅包括地理要素，而且包括这些要素的若干属性。这说明，在设计专题地图时，必须考虑地理特征的空间维度，并且了解每一变量的测量级别。把握住这两项因素即可把地图制作得更有视觉吸引力，更便于准确理解。

A5.16. 地理特征在地理数据库内用几何基本元素——点、线和面积——表示。对于一些不常使用的地图，还需增加维度，考虑体积和时空的概念。有时，采用哪一种几何形状表示真实世界的地理特征取决于地图或数据集的空间比例。例如，一个村庄或小镇在大比例尺地图上可以表示为一块面积，但在小比例尺的省级或全国地图上只可能是一个点(见图A5.4)。一条道路在省级图上是单线，到了城市地图上则变成了双线(即具有了面积特征)。

A5.17. 重要的是要记住，边界和位置并不总像一幅地图或地理数据库里那样清楚地界定。对于复杂的真实世界的地理特征，常常需要进行泛化、简化或抽象化处理，才能将其纳入计算机数据库。例如，很多真实世界的地理特征



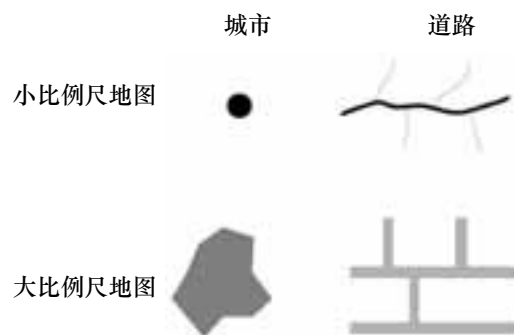
图A5.3  
人口密度专题地图实例



并没有清晰的边界。林区与非林区之间经常有一个过渡区存在。如果把森林表示成具有面积特征的物体(而不是用点来代表每一棵树)，肯定会丢失一定的信息(见图A5.5)。

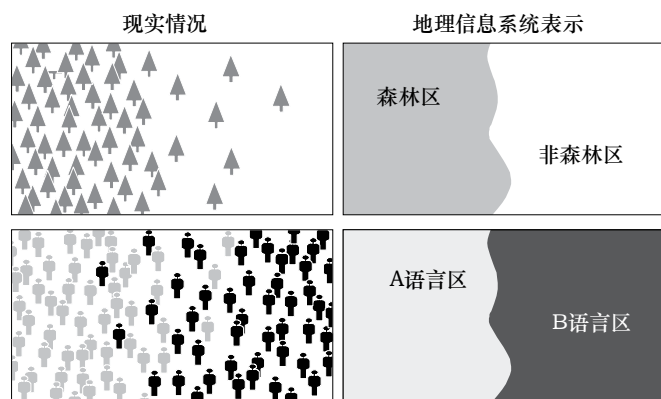
A5.18. 在社会经济领域也是如此，比如，在种族或语言群体分布方面也会遇到边界定义不当的情况。尽管有时候，这类群体具有非常明确的分布格局，但每一特定地区的边缘地带总会存在不同种族或语言群体杂居的情况。绘图人员有时使用斜线涂盖的方法代表此类难以界定的边界，但在地图上仍无法解决要把边界画在何处的问题。

图A5.4  
在显示空间特征时泛化的效果



图A5.5

为了便于地理信息系统进行表达，复杂的现实状况有时需要简化



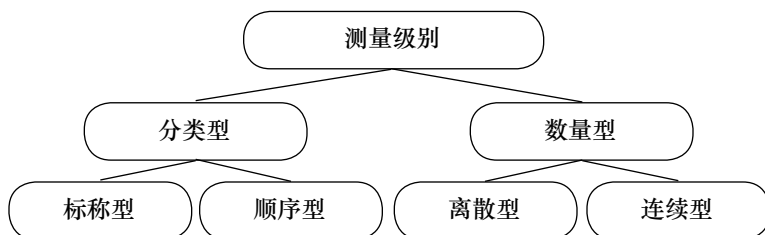
(b) 测量级别

A5. 19. 同样重要的问题还有如何计量地图绘制过程中所需变量的问题。主要区别是分类型与数量型信息的区别(见图A5.6)。分类型数据又可分为标称数据和顺序数据。标称数据，亦称定性式数据，直接描述地理特征的类型，但在类型之间不存在自然的顺序。例如房屋的类型包括石制或木制框架结构。而顺序数据则意味着可以排列出档次和级别，尽管档次或级别的差距并不明确，但仍能按顺序排列。例如，基于调查结果，可以把人们的生活水平分为低、中、高三等级，但我们可能说不出低级和中级之间的差别是否与中级和高级之间的差别一样。

A5. 20. 如果能对各种类型之间的差别加以量化，即可获得数量型数据。离散型数据可以清点，例如清点每个家庭的卧室数目或总人数。连续型或比率型变量可以取得任何希望的数值，因此能达到很高的精确度。对于人口普查数据，连续变量通常是加总人口普查单元计算的各种指标，如人口密度、拥有安全饮用水的人口比例或总的生育率等。

图A5.6

变量的测量

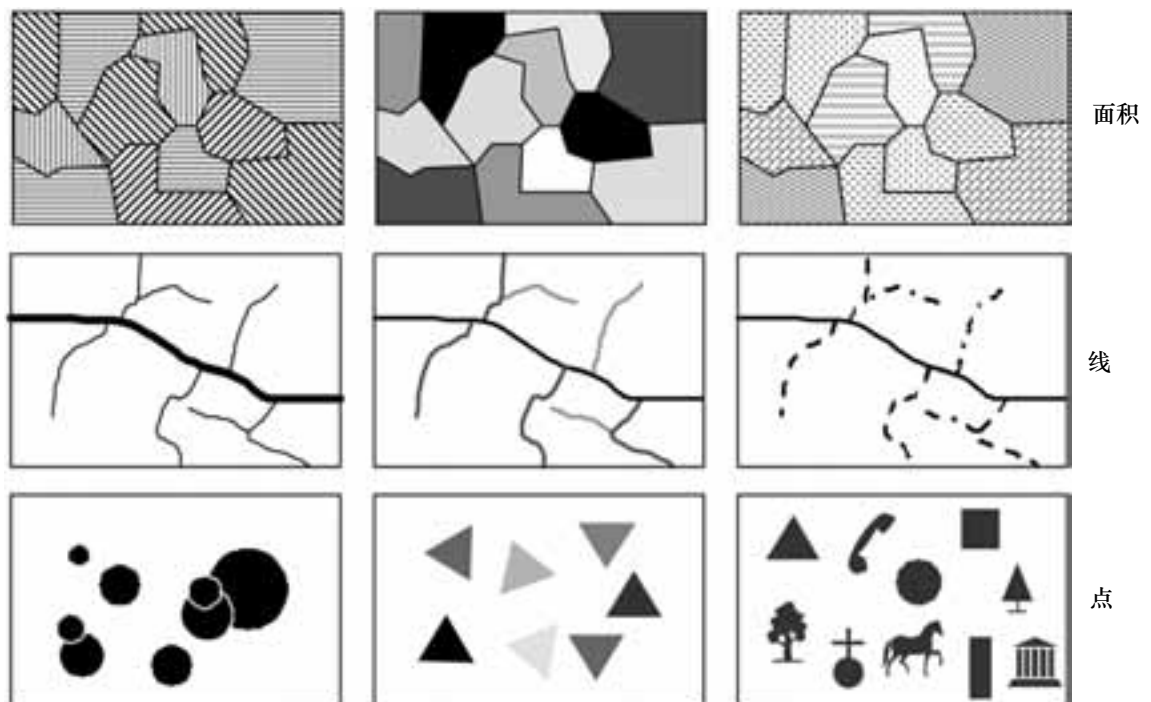


## (c) 图形变量

A5.21. 在专题地图上，图形符号向观察者展示出地理特征在数值或类型上存在的差别。地图绘制使用的符号概念与Bertin针对图形设计开发的符号概念基本相同(1983年；另见MacEachren, 1995年)。Bertin说明了以下各种图形变量的区别：

- **大小**。大小是区别顺序或数量差别的标志。大小对于点状或线状特征最为重要，例如，为了显示乡镇或城市大小使用逐渐变化的圆圈，或为了显示地区间移民量的大小而使用宽度不等的线条或箭头。
- **方向**。方向用于表现，例如，面积特征的交叉阴影。此外，几何点的特征也可显示成各种方向。方向并不意味着变量规模大小的差别，因此显示标称数据时十分有用。
- **纹理**。纹理指的是在面积与面积之间表现出多种变化的连续图案，用来表示顺序或数量之间的差别。这是一种非常有用的简化方式，当输出装置无法打印彩色或灰度级别时，可用纹理进行补偿。纹理还可以用来表示两种变量叠加在一起的分层信息。当然，地图的清晰度可能因此受到损害，但对初步分析应用仍非常有用。
- **形状**。形状对于点状特征最为重要。在商业地理信息系统和桌面地图绘制软件中，有大量的符号集和字符集可供使用。有一些是地图绘制行业众所周知的代表公共建筑的符号，如寺院或医院等。

图A5.7  
表示多边形、线和点的图形变量



- **颜色。**颜色最适合表现数量差别和一定程度上的顺序差别。在地图设计过程中，颜色的选择是最重要的问题之一。对此，下文还要进行更详细的讨论。

A5.22. 原则上，这些方面的任何一项都可以应用到所有类型的地理特征，即点、线、面积之上。但是在大多数实例中，仅用图形变量的一个子集代表不同的地理类型。图A5.7给出了一些实例。专题地图选择的图形变量要与地图绘制过程中使用的指标测量类型保持一致。例如，在数值表达方面，最重要的是大小和颜色。用点状符号的形状或面积特征的纹理表示不同的标称值。

### 3. 专题地图的类型

#### (a) 离散式地理特征图的绘制

A5.23. 编制后发布的人口普查数据包括很多诸如地区或点查区一类的报告单位的数字加总。等值线图能够最好地从地理方面表现这类数据。“等值线(choropleth)”是从希腊字“choros”(地方)和“pleth”(值)演变而来的。等值线图用以表示离散报告单位的数据，而离散报告单位的确定常常与数据的实际空间分布无关(如行政区边界)。用来表示每一报告单位的符号(颜色和图案)取决于单位数值。等值线图不同于所谓的“面积式”地图，报告单位在后者的情况下取决于数据，例如，在一幅显示森林覆盖率的地图上，报告单位根据林区和非林区的边界划分。

A5.24. 前面图A5.3中已经有一个等值线图的实例。建立等值线图的第一步是把报告单位数据值的全部范围分成一套不同的类别。每一类别指定一种颜色或深浅不同的图案。鉴于点查数据所具有的自然顺序，因此在颜色和色调深浅的选择方面有一定的逻辑可循，例如颜色从浅到深或图案从稀疏到致密。这样做的目的是让用户对每一报告单位值的大小有一种直观感。用于涂绘等值线图的符号可通过多种方式确定，这取决于变量的类型、数据值的范围以及打印地图的输出装置。符号选择非常重要，因此下文还要详细讨论。

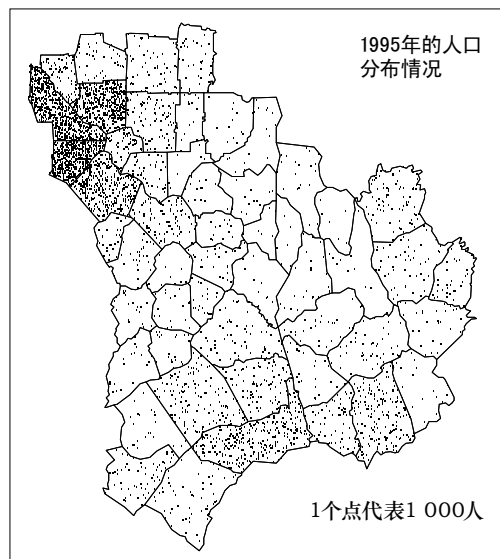
A5.25. 等值线图擅长表现数据值在地图上的总体分布，以及比较不同地图上的分布情况。一般不可能获得每个报告单位的准确数值，因为颜色和灰度仅能表示类似值的范围。准确信息最好能以数据表格的形式表达，或者在地理信息系统内通过交互式查询获得。

A5.26. 用于制作等值线图的数值几乎总是以比率、比例或密度的形式出现。它可以是与地理有关的比率，例如通过用人口数除以面积得出的人口密度；也可以是一般性比率，分母不是面积而是一个数值，例如人口出生率就是指每千人的大致出生数目。在大多数情况下，我们根据某一社会经济变量绘制地图时，报告单位的大小并非完全一样。例如，地区或省份在面积和人口方面经常存在相当巨大的差异。如果根据计数变量(例如人口总数而非人口比率)绘制地图，那么最大的地区将使用最深的颜色，即使其人口少于其他地区时也是如此。因此，不宜按绝对值绘制等值线图。

A5.27. 还有一种说明计数数据的方法是“点式地图”。1830年，法国在绘制全国人口分布图时首先使用了点式地图。在点式地图上，一个符号点代表一个或多个变量单位。例如，每个点可以代表1 000人或1 000个家庭。通过这种方法，把变量数值以密度不同的小圆点表示在报告单位上。图A5.8就是一个典型点状地图，表示人口分布状况。

A5.28. 如何放置这些点有两种方法。制图人员可以基于每一地区实际人口的分布情况选择点的位置。例如人口稠密的市区或市区周围放置很多的点，而人口较少的农村则放置较少的点。有时可以根据土地使用情况确定每个报告单位的点密度。还可以使用实际遮盖法，在无人居住的地带，如水体、繁茂的森林或自然保护区不使用点。

图A5.8  
点状密度地图



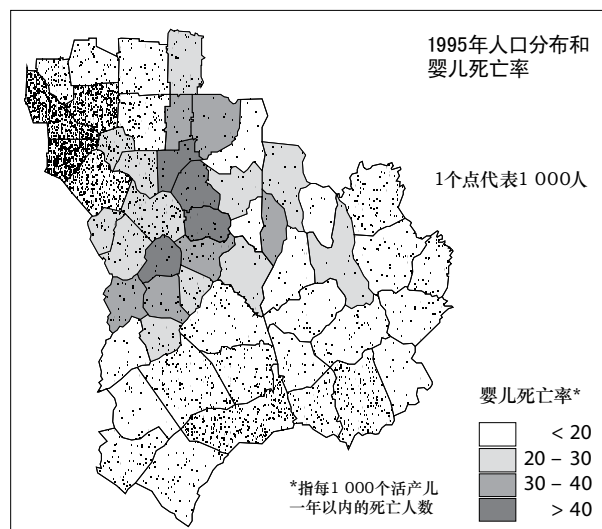
A5.29. 另外一种方法是每个区域内随机地放置点。这种情况下，点的密度反映的仅是总值密度。地理信息系统和桌面地图绘制软件提供的点密度功能通常为随机置点方法。用户只能控制点的大小以及采用何种符号表示一个点。尽管简单的圆点最清楚，但人们还是经常愿意挑选一个能代表变量特性的符号。

A5.30. 有一些专门的软件可以借助其他数据层帮助布点。手工布点可以结合制图人员关于变量分布的了解进行，但过于麻烦。

A5.31. 点状地图是一种表达密度信息的有效方式，只要布点方式遵循地图变量的实际地理分布状况，或在每一报告区内均匀分布即可。这种方法有一个非常大的优点，即复印或打印的效果非常好，因为它们都是单色黑白图。点密度图可与等值线图一起使用，同时代表两个变量，见图A5.9的实例。图中的实

图A5.9

## 点密度图与等值线图的组合



例说明，高人口密度与高婴儿死亡率之间并无关联。在这种情况下，点的密度不要太高，防止区域内的颜色或灰度不易辨认。

## (b) 标称点数据

A5.32. 点状图中最简单的情况就是每个点代表一个离散元素，例如一个农场或一个医院。这种标称点数据代表的是特征物的类别而不是计数或尺寸大小的属性。在简单的点符号图中，点的位置能正确地反映出地理元素的所在位置。利用大小、颜色或符号可以说明不同类型的特征物，如图A5.10所示的保健中心与医院的分布情况。也可以用简单的几何图案代表不同类型的点状特征物，如圆圈、方块或三角。有些桌面地图绘制软件或地理信息系统软件允许用户自己为地图的特征物指定一个相配的符号。如图A5.10中的地图表示的是两种健康设施的分布情况，为了便于理解而使用了不同的符号。符号一般为字母或位图，大多数软件都有自己的字库，其中含有很多不同的地图符号，如交通运输、公用事业或设施等。有些系统还允许用户导入自己设计的位图符号。不过，最好的做法是尽可能使用标准符号集，以减少复制时的特别要求。

## (c) 比例点符号

A5.33. 点状符号也可以用于代表特定地点中的数量。例如，常见的人口普查地图可以显示主要城市的位置及大小，所用的圆圈或方块的尺寸与该处的地理特征的数值的大小成比例。这种地图称为比例符号地图或者分级符号地图。这种地图适于表现某种变量的绝对值，不适于表现相对值，如密度、比率等。

A5.34. 有两种类型的分级符号地图。第一种情况是，用数据指代城市或家庭等点状特征物。符号的位置与特征物的位置相对应(见图A5.11A)。第二种情况是，用符号代表面积特征物的数值，如地区。在这种情况下，需要在每个

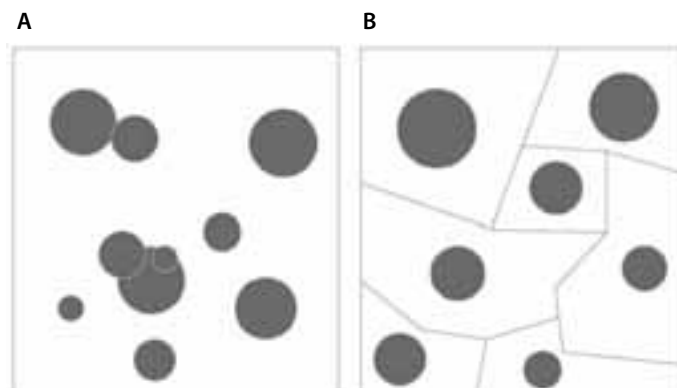
图A5.10  
绘制离散点对象图



报告区内选定一个代表地点(见图A5.11B)。需要注意的是,大多数系统都会在每个圆圈之外画一个晕圈,以便区分靠得很近的圆圈。系统进行操作时,先画出最大的圈,以防覆盖住小的圆圈。

A5.35. 如上所述,计算机制图软件允许用户选择一个能反映地图专题的符号。这种形象化的符号可以使地图看起来更有趣。然而,过于复杂的符号会分散人的注意力,忽略地图表达的主要信息:不同区域的变量相对值。我们可以比较一下同一专题地图的两种不同版本,表示的都是电话数量,见图A5.12。尽管左图的电话符号相当简单,但很难判断变量的大小,因为左图不如右图简单。制图人员常常要掌握好一个尺度,既要使信息容易理解,又要有吸引力,必须在两者之间寻找平衡。从几乎所有的例子来看,比较好的效果采用的都是简单符号,使读者的注意力集中在所研究的变量大小之上。

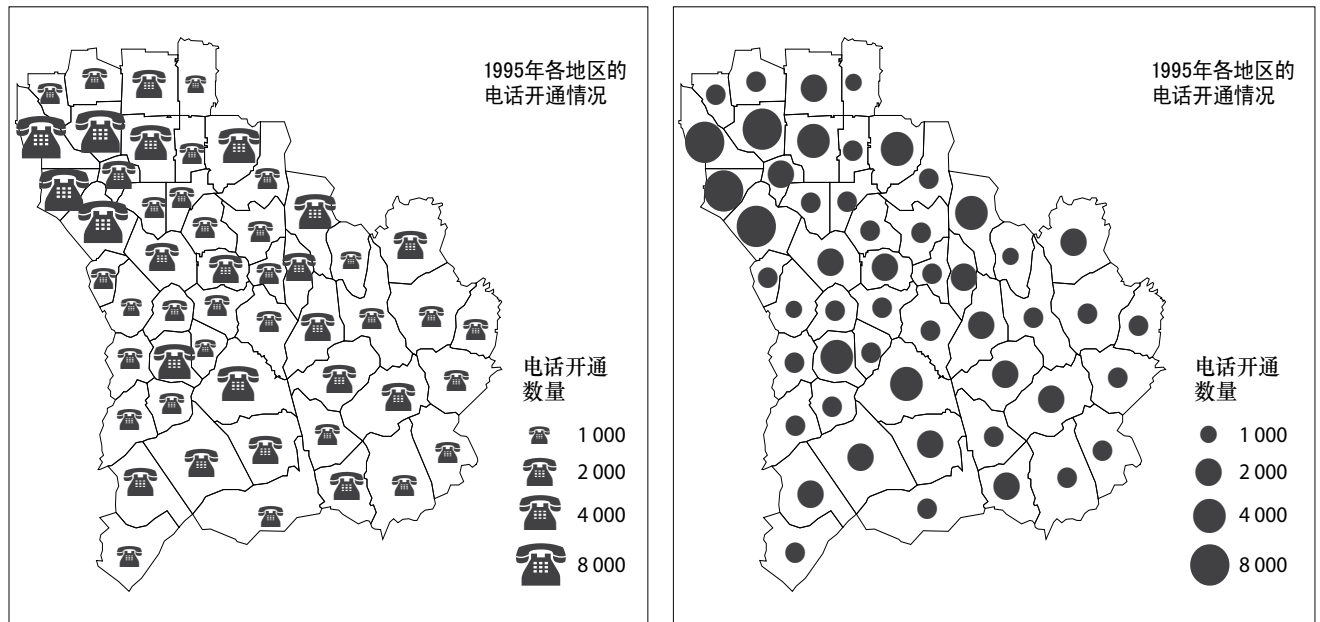
图A5.11  
用于点状和面积特征物的比例符号



A5.36. 比例符号也可以同时代表两个变量。例如，圆圈的大小可以代表报告区内的住户数量，颜色或灰度代表安装电话的住户百分比。再强调一次，绘图人员一定要避免地图上信息过多。如果报告单位的数量非常多或单位非常小，最好把两个变量分别显示在不同的地图上。

图A5.11

图画式符号与简单图形符号的对比



A5.37. 除了圆圈之外，其他常用几何符号还有方块和三角形。通过改变三角形的方向，可表示具有方向性的变量的大小，例如，每个报告单位的迁入或迁出人口数(见图A5.13)。如果配以不同颜色或灰度则更容易辨认。

A5.38. 与分级符号地图有关的还有一类地图，其中的数值差用标准化符号的数量表示，每一个地理单元均画有这种标准化符号。例如，总人口数可以按图A5.14的样子表示。在专题地图的绘制中，这种地图过去十分常见。但要注意，图画式符号很容易使画面变得杂乱不堪和难以解读。当有大量不同信息值存在时，使用比例符号会取得较好的效果。

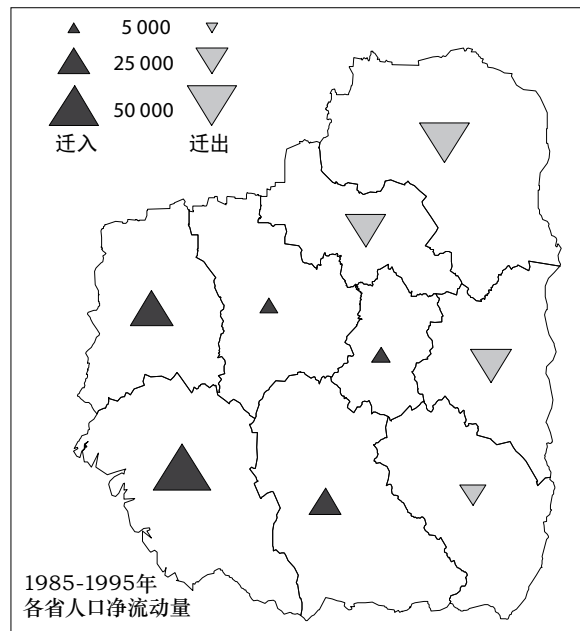
(d) 图表式地图

A5.39. 由于商业桌面制图软件和地理信息系统所具有的图表调用功能，在显示每种地理特征的统计信息方面经常采用图表方式。和前面讨论过的几种地图类型一样，图表式地图很容易出现信息过度拥挤的现象。遗憾的是，很多出版的地图已经出现了这样的实例，读者很难或无法从中读取有用的信息。

A5.40. 最常见的图表式地图有饼图、竖图或柱图式地图。以饼图式地图为例，饼图通常是按比例绘制的，它的尺寸代表分母的大小。图A5.15中的实例显示的是各主要宗教群体的地理分布比例。饼图的大小根据总人口数而变化。因此，图例中需要给出两类信息：用颜色代表每一宗教群体；用饼图的大小代表人口总数。



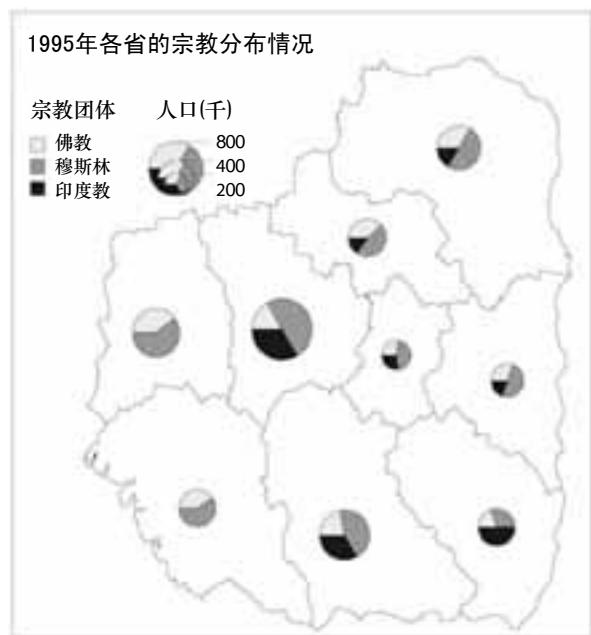
图A5.13  
用简单图形符号表示流动的规模 and 方向



图A5.14  
通过不同的地图符号数量表示每一种特征的数据值



图A5.15  
饼图式地图

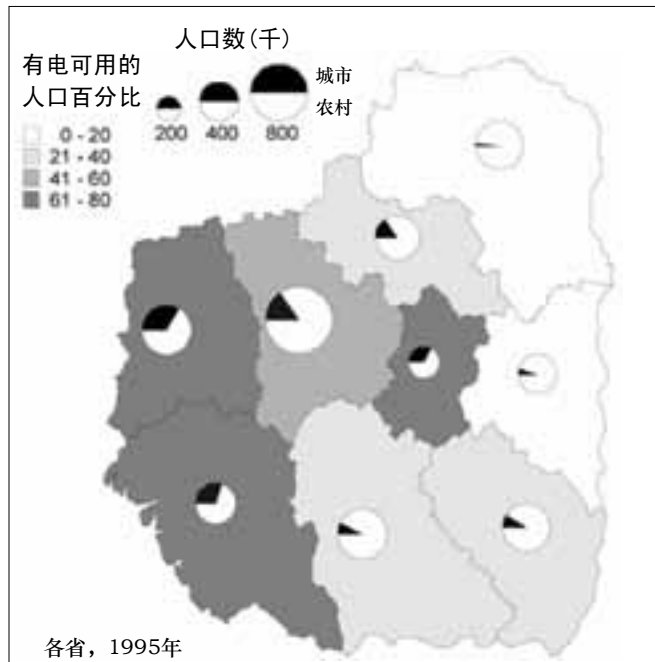


A5. 41. 如果地理观察值较小，并且只有非常少的人口群体，图表式地图可以取得最佳效果。例如，只包括两种类别的饼图可以和简单的等值线图一起说明几个变量(见图A5.16)：供电能力的不同分布情况、各省的人口总数以及城乡人口比例。从图中可以看出，那些城市人口多的省份，其使用电力的人口数量也多。设计良好的地图，不仅不会出现符号、颜色以及灰度过度拥挤的现象，而且能进行多种变量分析。然而，饼图式地图及类似地图很容易出现不易理解的情况，因此只限于在没有过多符号和信息类别的情况下阐述地图信息。

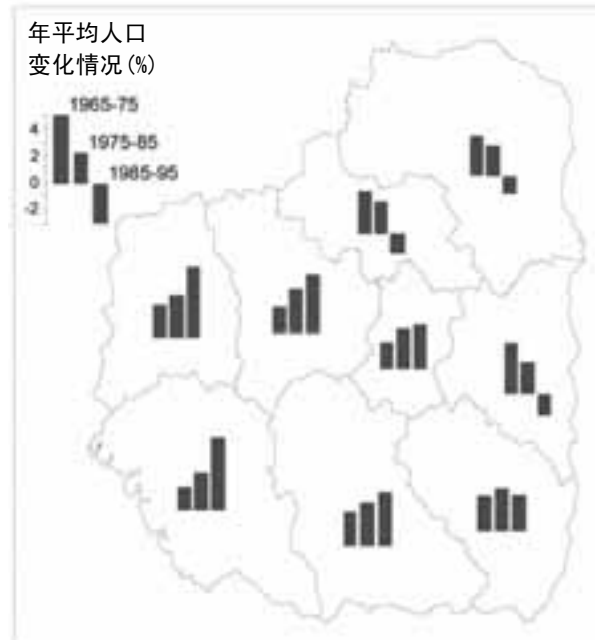
A5. 42. 图表式地图还可以表现随时间变化的各种趋势。以图A5. 17中的地图为例，该地图展现的是过去三次人口普查发现的各省人口变化的平均年度百分比。地图中的竖图非常简单，既没有边框也没有基线，人口的增减一目了然。正如以前所述，我们需要了解的是随时间发生的变化趋势，而非精确的数值，而表达精确的数值最好使用表格。

A5. 43. 人口金字塔是另外一种非常适合表现人口普查数据的图表式地图。这种金字塔式地图可与报告单位的底图配合使用，表明国家各地区人口的年龄与性别的分布情况(见图A5.18)。人口金字塔是一种非常复杂的图表式地图。这意味着只有地图中的地区数量不多时才能合理地使用。一般来说，只有第一个地区级别的人口普查地图才使用人口金字塔。另一个实际问题是，商业地理信息系统和桌面地图绘制软件不能自动生成金字塔式地图，必须通过外部建立，例如，首先通过电子表格软件建立，然后在图形软件中将其添加到底图上，或利用桌面制图程序的版面设计模块进行合成。

图A5.16  
等值线与饼图结合的地图

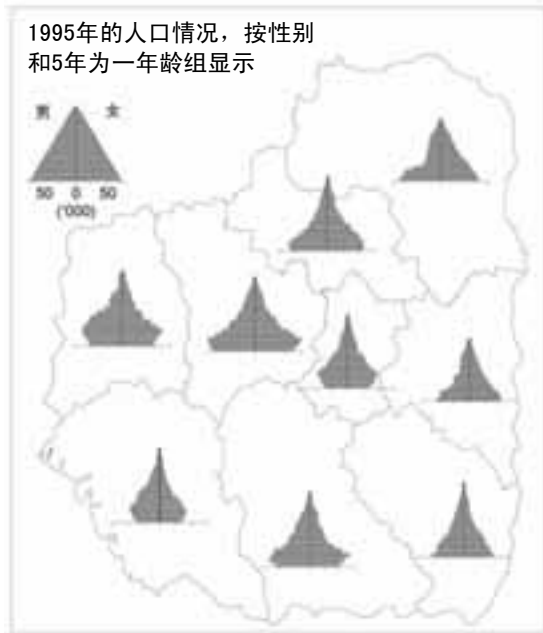


图A5.17  
通过直方图显示不同时期变化的地图



A5. 44. 如果使人口金字塔的形状发生一些变化, 即可用于显示各个地区的情况。如果各地区的年龄和性别分布都相当一致, 那么得出的地图就没有多大意义了。从图A5.18中能够看到, 东南省份的生育率在过去15年间下降了, 而北部省份则没有下降。再有, 在东北省份, 人口性别比率有偏离正常的迹象。

图A5.18  
地图与人口金字塔的组合



在那些有经济活力的年龄段的人口中，女性多于男性。在西南部，情况似乎正好相反。

A5. 45. 性别比率的变化可以用另外一种竖图加以说明，见图A5.19。该图显示出每一个省份男女人口的冗余和欠缺数。在人口金字塔地图中能够看到

图A5.19  
在地图上显示人口的性别比率



的趋势在这里表现得更加清晰。然而，这种地图相当复杂，看上去不是很有吸引力。关于性别比率的其他表现方法下面还要介绍。

#### (e) 流量地图

A5. 46. 移民是人口学中的一个变量，表示人口在一个国家内从一个地区向另一个地区的迁移(国内移民)，或从一个国家向另一个国家迁移(国际移民)。在地图上表示移民有很多方法。可以通过等值线图表示迁入、迁出和净移民比率。移民量可以用分级符号图表式(见图A5.13)。还有一个办法是，如果可以获得有关移民的完整信息，可以使用流量地图，亦称流线地图。这种地图可以表现移民的若干方面：移民流动的路线和方向(来自何处，去往何处)，可用箭头符号表示走向；以不同宽窄的线条表示移民流量等。

A5. 47. 移民地图很快就可以变得非常复杂，即使我们列举的实例中只有9个报告单位，但流量线就可能有72种，这还不包括国际移民和省内移民。因此，各种流动路线一条不少的整个地区或全国地图几乎没有。有几种变通方法。一种是忽略最小的移民流，只表现最大和最有意义的移民流。再一个办法是分省制图，仅表现迁入或迁出一省的移民流动情况(见图A5.20)。就这个例子而言，一共需要画出九对地图。即使一幅简单的地图也会搞得拥挤不堪。如果起始地和目的地相距遥远，制图员还得画出各种蜿蜒曲折的弧线。

A5. 48. 在使用箭头符号的流量地图中，人们会受箭头宽度和长度的引导。长而细的箭头看起来比短而粗的箭头更醒目，因为其面积更大。尽管绘图人员有时希望利用这一事实来指出来自遥远地区的一个令人感兴趣的移民流动情况，但读者往往难以根据长短不一的箭头评估流量的相对规模。如果侧重于来自各个地区的移民绝对数量，还有其他更合适的方法。例如，通过表示起始地或目标地的分级符号来说明移民流动的数量(见图A5.20)。

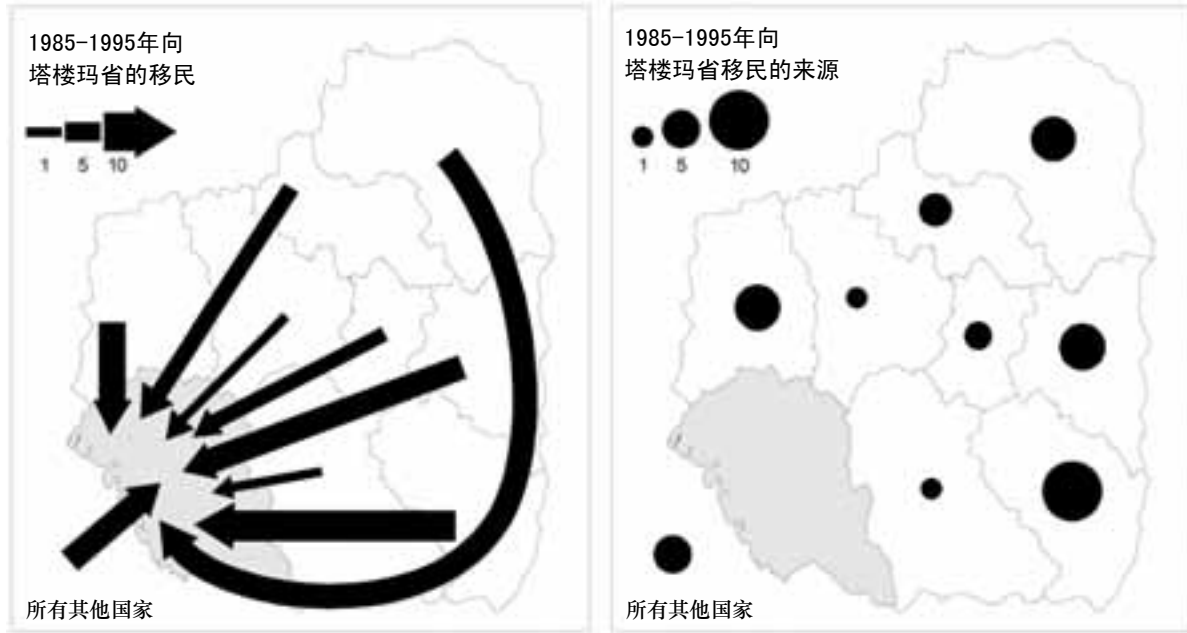
A5. 49. 采用特殊类型的分级符号，一张地图可以同时显示每个省迁入和迁出的移民，如图A5. 21所示。不同颜色或不同灰度的半圆用以区分迁入或迁出。

#### (f) 连续现象地图

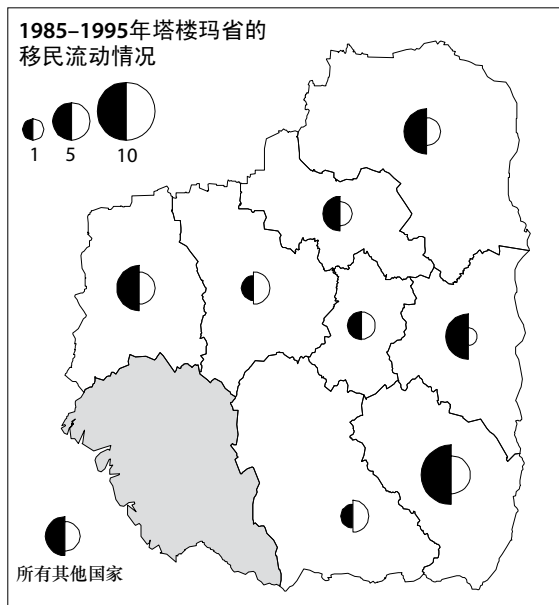
A5. 50. 上述各类地图都适用于以离散地理特征为基准的数据，例如地点和区域。但有些地理现象是连续的，例如，气温和标高线的数字就是如此，它们沿空间平滑变化。而且，我们还可以把人口分布或多或少地看作是一种连续变量。报告区比较随机，而且这些单位的表格化加总值把空间变化隐藏到每个单位之中。人口地图和地理信息系统数据集有时也把人口密度和分布情况显示为连续性的变化，并且这种现象正在日益增多。

A5. 51. 纸张地图或计算机数据库几乎无法表示真正的连续性。即使我们能从理论上为国家的每一个具体点推导出不同的数值，为了绘制地图，我们需要以某种方式对这些数据进行分类。图A5.22说明了几种处理方式。

图A5.20  
各种表示地区间移民流量的方法



图A5.21  
移民迁入和迁出的表示法

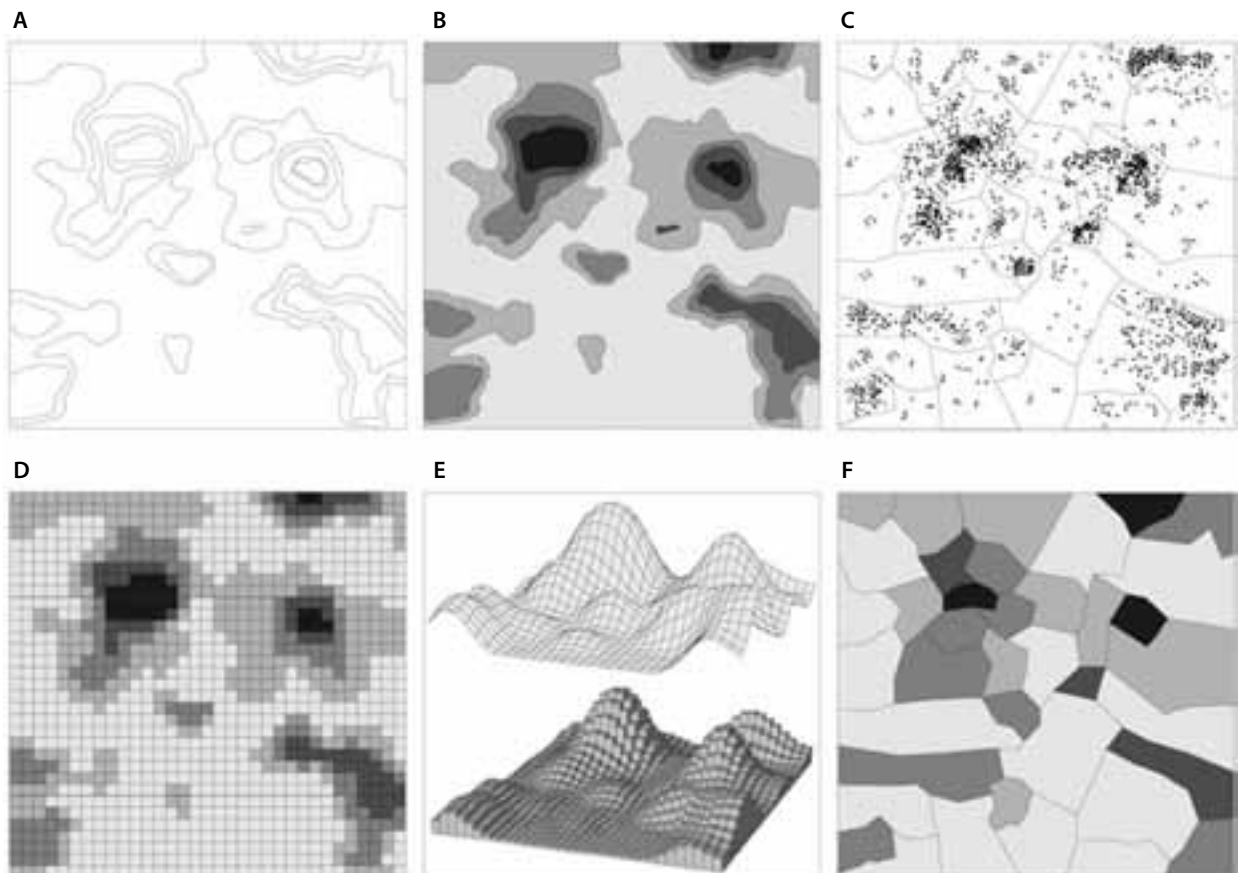


A5. 52. 表现连续数据最常用的方法是借助等值线或规则光栅网格。等值线(Isolines)一词起源于希腊字“iso”，意思是相等，它是连续数值构成的线段，有时也称为等高线(见图A5.22A)。它们在地形图上用于表示高度。轮廓线地图也可以涂成不同的灰度，使它们看起来更像等值线图(见图A5.22B)。颜

色可以用来表示两条等高线之间的数据范围。点状式地图也可以用来提供更连续的人口分布视野或类似的变量形象。正如前面所述，大多数地理信息系统软件是通过在每个报告单位之内随机置点的方法生成点状式地图的。在这种情况下，点状式地图与等值线地图相比，并没有让人获得更多信息。但如果这些点是按照土地覆盖状况或村庄的位置放置的，则有可能获得更为连续的变量分布图像(见图A5.22C)。

A5. 53. 为了进行模型化处理和分析，连续数据在地理信息系统内通常存储为规则的光栅网格(见图A5.22D)。选择网格单元格的大小，使其足以保持数据集内的变量特征，但细腻的网络将使文件尺寸变得非常大。最后，计算机地图绘制软件以及一般图形软件有许多方法能把连续变化的数据集表示成一个表面。图A5.22E中的两个例子说明的是网状框架模型和二维竖图。这类技巧在表现基于数字化标高模型的地形信息方面非常有用。有时，用这类图说明人口分布情况很不错。在这类地图中，小山和山峰代表高密度人口聚集区，而山谷则说明该处人口稀疏。对于人口以及类似的社会经济信息，经常难以获得表面各处真实的空间分布情况。人们能很直观地理解标高的高度，但很难把表面高度与各自的变量值联系在一起。因此，一般来说，更适合的做法是多使用一些标

图A5.22  
各种表示连续性数据的地图绘制方法



准地图绘制技巧。为了便于比较，图A5.22F提供了一幅报告单位不是根据数据分布情况确定的等值线地图。

### C. 数据分类

A5.54. 以上各节讨论了绘图人员在地图上表示专题信息时使用的工具。地图设计者必须选择图形变量以及与该变量相适应的专题地图类型。有些实例中，符号类型与变量值之间有一对一的匹配关系。例如，在仅仅表示少量的标称类型时就属于这种情况，其中点状符号的大小近似但形状不同。然而，即使是分类数据，也经常需要用相同的图形符号表示几种具有类似数值的地理特征。例如，单身家庭和多人人口家庭都可以用同样的点状符号表示。数值型的数据几乎毫无例外地需要进行分类处理，然后才能根据符号尺寸和颜色进行匹配。

A5.55. 将数值接近的观测值归并到一起并以相同的图形符号表示的过程叫做分类。这个过程类似于统计工作的分类，做法是将各种数值分组，使同一种类的观测值尽量相同，不同种类的观测值尽量不同。计算机地图绘制软件提供了一些默认方法，将符号分配给数值或数值范围。尽管图形用户界面有了改进，默认方法也可能不适合要绘制的变量——大多数情况是不适合的。事实上，自动分类工具经常导致不恰当的情况出现，甚至对地图设计造成误导。以下段落将对一些分类方法进行较为详细的探讨。

A5.56. 数值型数据类别通常是相互邻接的数值范围。类别的数目取决于几个因素：数据分布情况（即数据集内数值的变化）、数据表达的预期准确度，以及输出设备分辨细小颜色和纹理差别的能力。使用更多的类别不一定能对专题地图起到改善作用，因为它增加了观察者分辨各种类别的难度。更重要的是确定类别范畴，以便准确反映数据集的变化。

A5.57. 究竟哪种分类技巧更适用取决于变量的数据分布情况。既要保持准确、又要使统一分布的地图数据集（即高、中、低三种数值的数量基本相等）具有视觉吸引力的方法对于有很多低值和少数高值的、反差极大的数据分布不会取得良好效果。

A5.58. 为了制作符合出版质量的地图，数据应该根据统计图表进行评估。遗憾的是，一些地理信息系统和桌面地图绘制软件的统计图表制作能力有限。不过，它们可以导出数据，供图表功能丰富的电子表格软件或统计软件使用。

A5.59. 在确定分类范围时，一种最有用的图表是等级序列图。所有的数据点根据数值从低到高排列，依次绘制，x轴表示每一个观测值的等级，y轴表示数据的数值。相邻数据点之间的垂直间隙或“自然中断”是良好的分类界限，但和希望得到的类别数量相比，经常出现过多或过少的间隙。



A5. 60. 下面提供了一些常用的分类方法实例，其中的三个变量具有不同的统计数据分布。人口密度变量分布得非常不平衡。其中很多小数值的范围在每平方公里21到大约110人之间，并且只有几个非常高的数值。最大的数值(791)几乎是次高值(320)的两倍半。这种现象在人口密度方面十分常见。例如，高密度区可能是农业省份的首府。第二个变量是该地区的识字率，这一数值分布相当均匀，从等级序列图的观测值线段上看，几乎是直线，没有极端的数值。

A5. 61. 第三个变量实例是总生育率。从等级尺寸图上看，在最低的观测数值上出现激增，中间一大段激增减少，在接近右侧的地方，再次出现高位观测值激增的现象。这说明了一种正常分布，特点是过高或过低的数值均比较少，众多的观测值处于中间范围。当然，这些例子只是为了起到说明目的。在其他地理区域，同样的变量可能会有非常不同的分布状况。

A5. 62. 这些例子说明，一幅地图的外观在很大程度上取决于分类方法，这些方法可能适合于特定的数据分布，也可能不适合。这就证实了一点：在使用地理信息系统软件提供的自动分类功能时要谨慎小心。

## 1. 序列数据分类

A5. 63. 最简单的分类方法之一就是把数据值范围分成“相等的间距”(见图A5.23)。绘图人员首先确定出需要使用多少级别。数据值的范围(用最高值减去最低值)，除以种类数量得出增量，增量也叫公差。第一级的范围是从最低值到最低值加增量，随后的各级是在上一级高限范围值上加一个增量。图例中的低精度数字有时需要四舍五入。

A5. 64. 对于人口密度变量而言，最低值为21，最高值为791，因此范围为770。由于我们需要使用五个种类，因此公差为770除以5，得154。第一级从21到175，下一级从176到329，依次类推。

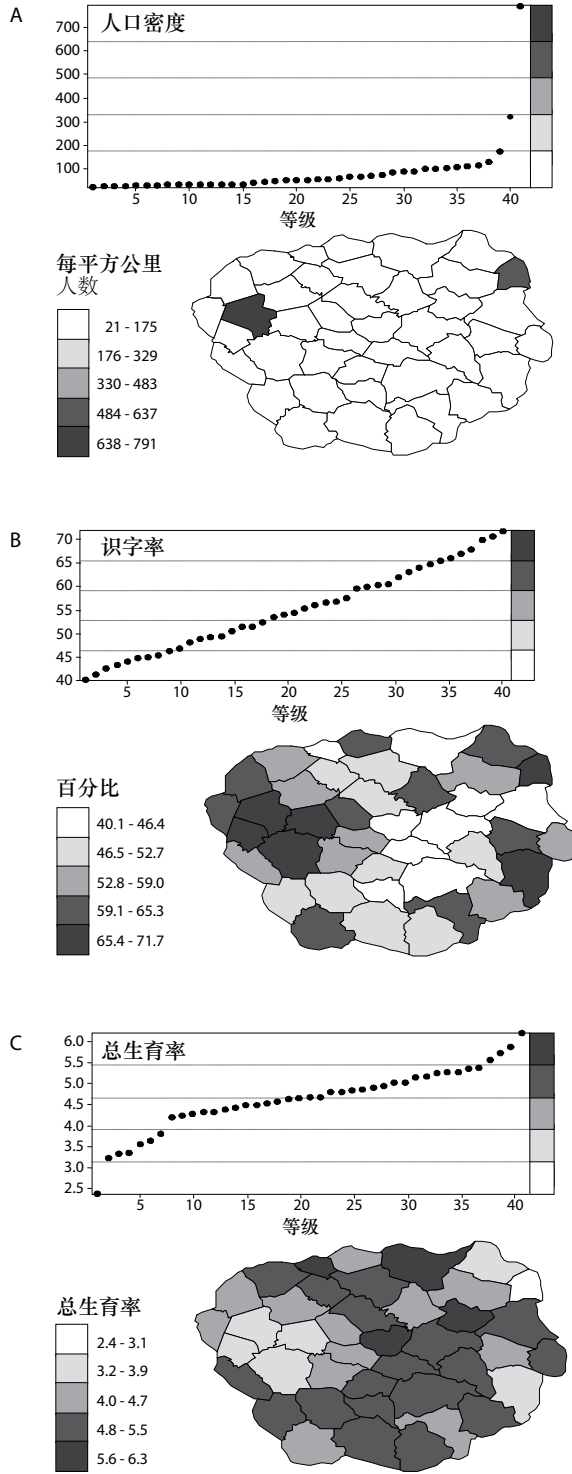
A5. 65. 这幅人口密度地图说明了问题出现的原因。数值范围受一个非常大的数值的影响。公差过大造成第一级范围几乎包括了除两个观测值外的所有观测值。显而易见，制作这种地图没有多大意义。

A5. 66. 然而，这个方法用于表现均匀分布的识字率则能取得比较好的效果。数据集划分后，每一级包含大致相等的观测数量，由此产生的地图能够很好地表现各地区的识字率分布情况。

A5. 67. 最后是总生育率地图。它的问题与人口密度图相似，但没有那么严重。只有一个观测值在最低一级的范围内，整个图被中间级别范围的数值所占据。然而凑巧的是，第二和第三级、第四和第五级之间出现中断，正确记录了在数据分布中也出现的中断。

A5. 68. 对于序列数据分类来说，除了等距法之外还有其他的分类方法。有一种分类方法使用了几何递增原理，如0-2、2-4、4-8、8-16等等。这种方法适合于人口密度变量那样的高反差数据分布。

图A5.23  
相等的间距



## 2. 统计分类法

A5.69. 这种分类法使每个类别中含有大致相等的地理观测值，借助于分位点的统计概念实现，即将数据集分成同样数量的观测值类别。如果有四个类别，则称为四分类，如果有五个类别，则称为五分类，依次类推。

A5.70. 为了确定分位点，用观测值的数目除以预定的类别数目。必要时，通过四舍五入得出最接近的整数。在等级序列上，第一批“n”个观测值归入第一类，第二批“n”个观测值归入第二类，依次类推。所有余数均指定给第一或最后一个类别。

A5.71. 分位点制图技术在很多桌面地图软件包中均可实现，因此在制图中十分流行。

A5.72. 三个样图看上去都不错。从定义上看，各种观测值分布良好，因此所有地图均可充分利用所有灰度级。

A5.73. 我们来看一下数据分布情况，对于识字率变量来说，这种分类法似乎非常合适。事实上，这张地图看起来与等距法地图没有多大区别。

A5.74. 然而，在人口密度图和总生育率图中，我们可以看出，这种方法把类似的数值归入了不同的类别。以总生育率为例，在最低数据组内(2.4-4.2)，有两个最高观测值与第二类的观测值比较相似。更有甚者，数值同样为5.3的三个观测值，一个归入到第四类范围，而另外两个归入到第五类(有些桌面地图绘制软件放松了数量观测值的判断标准，以避免出现这种情况)。

A5.75. 因此，使用分位点地图时也应当谨慎。常见的情况是，类似的数值归入到不同种类，不相似的数值却归入到同一种类。地图看上去虽然吸引人，但这种印象可能造成误导。

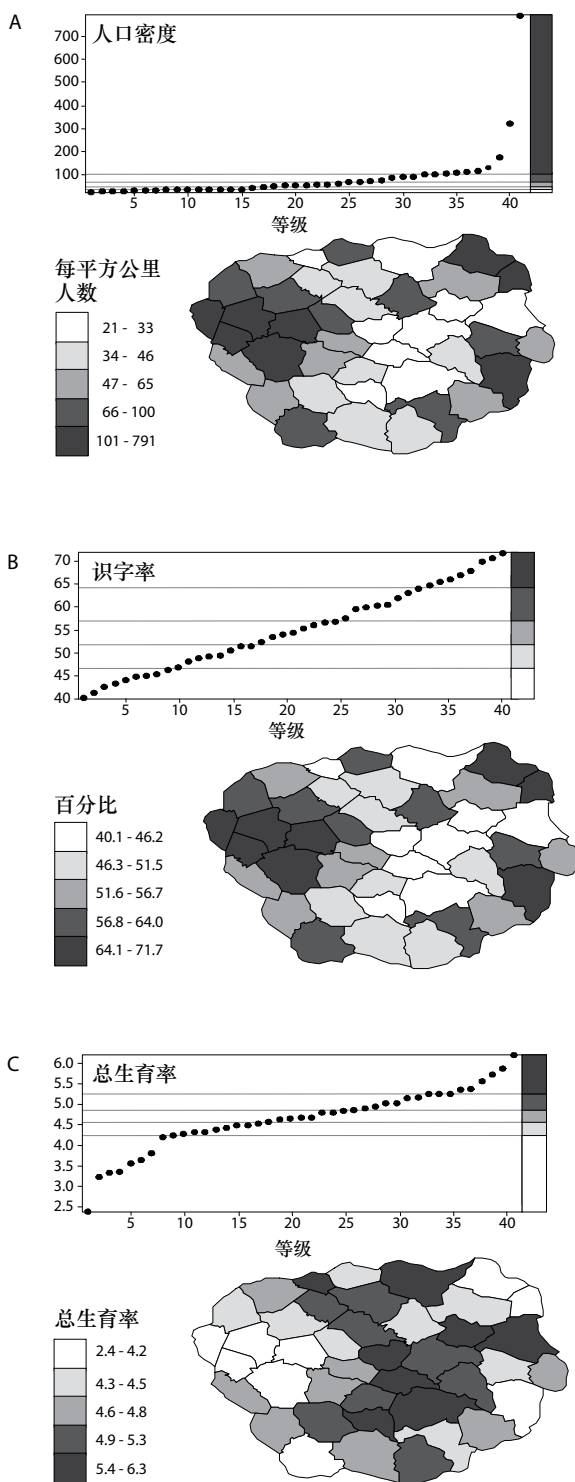
A5.76. 另一种统计分类技术基于数据分布的概括性度量。一种方法是通过变量分布的标准偏差确定类别的范围。标准偏差根据方差平方根计算。方差根据数据值和总体平均值之间的平方差的平均值计算。例如，对于识字率这一变量，标准偏差值为8.9。

A5.77. 地图类别以标准偏差为基础，因此可以显出具体的观测值(如地区观测值)与全省或全国平均值的对比。

A5.78. 类别在平均值上减去或增加标准偏差得出(识字率的平均值是55)。因此，级别范围的大小是常数，类似于等间距法。

A5.79. 对于识字率而言，第一个数据范围(40.1-46.2)与平均值相比，是大于一个标准偏差，但小于或等于两个标准偏差的数值。因为数据分布相当紧凑，所有的数值都在 $\pm$ 两个标准偏差范围之内，并且只需要四个类别。如图A5.25B所示，根据这种方法对识字率的数值进行分类，每个类别都有基本均匀的观测值数目，为地图带来良好的视觉对比。

图A5.24  
分位数(等频)制图



A5. 80. 可是, 这种方法对于人口密度变量来说则不太适合。因为小数值很多, 平均人口密度相当低(85.4), 而标准偏差却相当高(124.8)。第一类别与平均值相比在一个标准偏差之内, 因此实际范围在-39.5到85.4之间。另一方面, 最高值(791)又高出平均值五个标准偏差之上。这样就必须分出更多的类别, 而其中的大部分类别内一个观测值也没有, 所以不能这样划分。相反, 这儿呈现的地图的最大一个类别包括了高于平均值一个标准偏差的所有数值。显然, 对于这个变量, 标准偏差法不是良好的选择。

A5. 81. 对于表示总生育率的变量, 标准偏差法效果较好, 平均值为4.6, 标准偏差为0.8。然而, 只有非常低的数值2.4进入最低的类别, 低于平均值两个标准偏差以上。

A5. 82. 标准偏差分类法具有直观的特点, 因为它与描述性统计技巧密切相关。如果数据分布正常, 方差相对较低, 效果应该不错, 最多六个类别就可包括所有的数值。

A5. 83. 标准偏差法可用于表达数据集内不同类型的趋势(见图A5.26; 另见Dent, 1999年)。图A5.25中的实例采用了从浅到深的灰度等级。相对于图A5.26A中所示的类别而言, 这些图突出显示了人口密度、识字率和总生育率从低到高的数值递增。事实上, 这就是标准偏差分类法最罕见的应用。

A5. 84. 这种方法更常用于强调离散趋势。例如, 为了表示收入水平, 可能需要强调最贫困和最富裕的地区。在这种情况下, 我们把强烈的颜色和纹理指定给那些偏离平均值一个、两个乃至更多标准方差的地区, 而把相对较浅的灰度指定给数据分布的中间地带(见图A5.26B)。

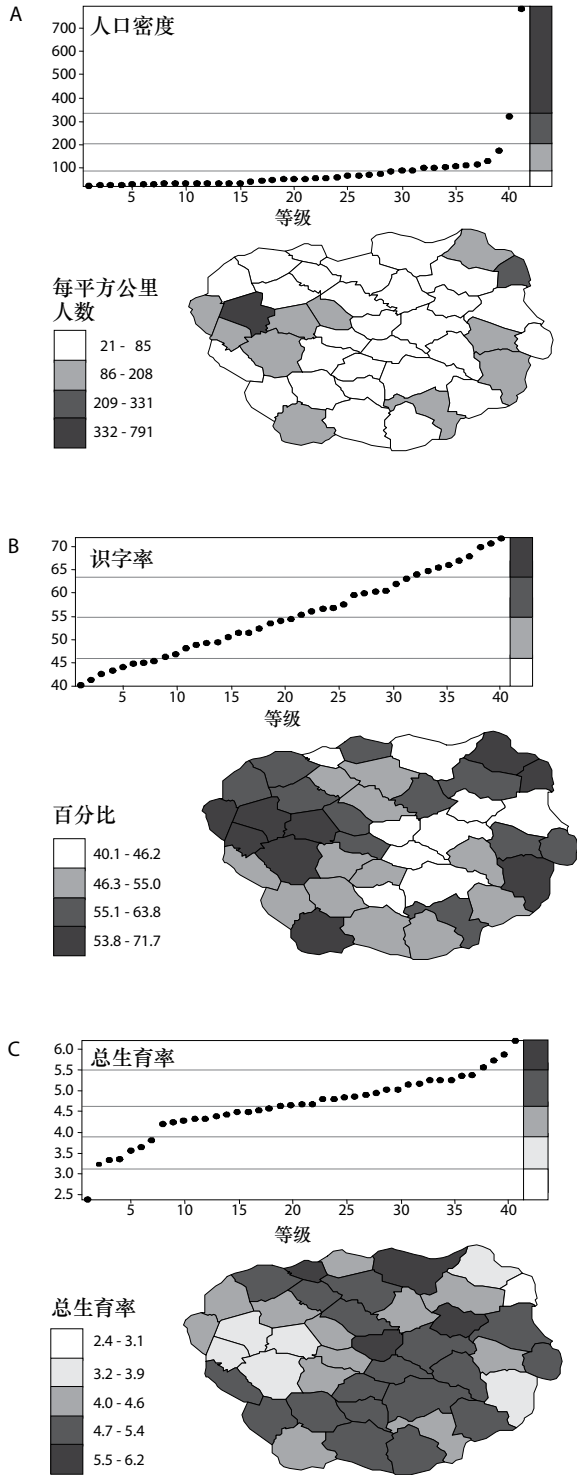
A5. 85. 如果只注意偏离平均值的距离, 不关心数值偏高还是偏低, 那么两侧可以使用同样的颜色。如果我们需要注意数值是在平均值之上还是之下, 两侧则要使用不同的颜色或纹理。例如, 在一幅彩色地图中, 低于平均值的类别可以指定为由浅到深的红色, 而超过平均值的地方则涂成相应的蓝色。

A5. 86. 在其他情况下, 我们需要强调中间范围(见图A5.26C)。例如, MacEachren(1994年)在Fothergill和Vincent(1985年)发表的对北爱尔兰(联合王国)地图研究的结果显示了新教徒与天主教徒的比例关系。在该地图上, 数值约为50%, 这说明新教徒与天主教徒的人数大致相等。在地图上, 中间等级以黄色强调, 在天主教或新教明显占据主流的地区则以较为柔和的颜色表示(分布为绿色和橘红色)。

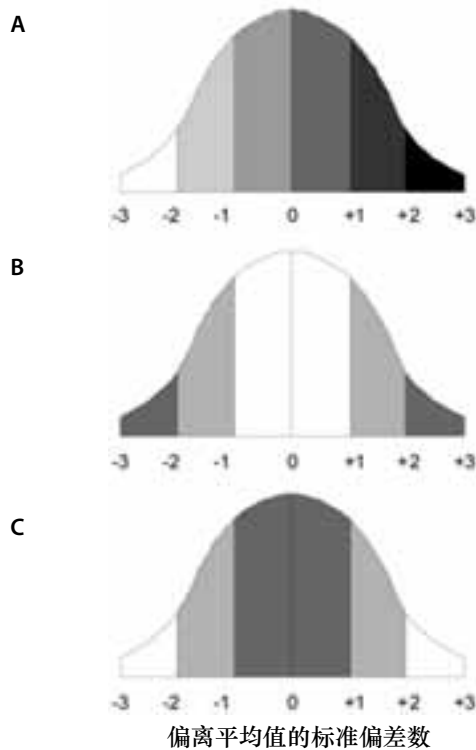
### 3. 自然中断

A5. 87. 正如前面几个例子中看到的那样, 对于非均匀分布的变量, 大多数方法制成的地图多少都存在一些误差。经常出现把类似数值归入不同的种类或把差别很大的数值归到一起的现象。从地图绘制的角度看, 合理的数据分类方法应当把每个类别内的数值差降低到最低程度, 而组与组之间差别最大。

图A5.25  
标准偏差



图A5.26  
根据标准偏差指定等级的灰度



A5. 88. 这个目标可以实现，方法是直接观察数据的分布情况，然后选择类别的中断点。这个方法的例子见图A5.27。对于总出生率变量，这种方法十分直接，从它的分布中可以看出几个明显的中断点。

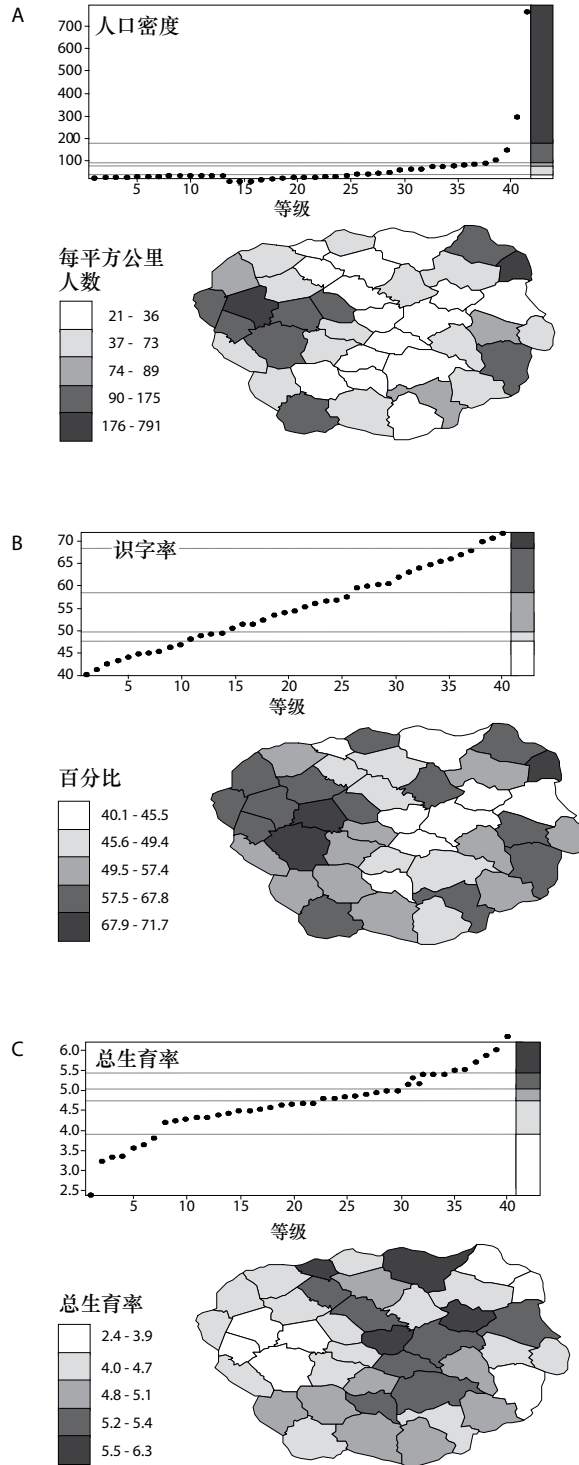
A5. 89. 然而，对其他两个变量则比较困难。对于人口密度来说，直接套用这种方法可能会把所有低值归入同一类别，而把较高的值放到不同的类别之中。因此需要调整，把低值范围的偏离值缩小，以求平衡。

A5. 90. 同样，对于均匀分布的识字率变量，类别中断不是非常明显，观测值之间的数值差变化不大。

A5. 91. 尽管如此，根据自然中断进行的分类充分考虑到数据的分布特点，因此，这种方法一般能准确地表示出地图的数据，并获得良好的视觉效果。

A5. 92. 除了依靠主观判断之外，还可以利用计算机确定自然或最佳中断点。许多地理信息系统和桌面制图软件提供这些功能，可以根据数据的分布情况自动确定自然中断点(Jenk优化分类法)。此外，还可以利用统计软件包的分类或集群功能。

图 A5.27  
自然中断点





#### 4. 无级别间隔的等值线图

A5.93. 所谓无级别间隔的等值线图并不要求制图人员选择分类方法。由于显示和打印技术的进步，计算机屏幕和打印机可以产生相当多的色调和灰度等级。对于没有类别或有 $n$ 个类别的地图来说，通过数据值本身即可直接确定出灰度等级百分比。例如，对于百分比变量来说，可以将0%(白色)到100%(黑色)的灰度等级与每一个观测值对应在一起。尽管制作或印刷时可以生成足够数量的灰度等级，但建议最好避免使用白色作为一种灰度或颜色，因为底面的背景就是白色。

A5.94. 实际上，这样做也可能得不到最佳效果。理由之一是很多变量的范围并不是从0到100，而是数值集中在一个较小的范围内。这样的话得到的地图可能只有非常浅或非常深的灰度。我们可以通过“拉伸”数据分布的方法克服这一问题，用最浅的颜色代表最低值，用最深的颜色代表最高值，这样制作的地图有利于读者阅读。

A5.95. 然而，一般来说，容易区别的灰度和颜色数量是有限的。连续色调方案仅为分析目的使用，一般的做法是将数据值分为不多的类别。

#### 5. 外部数据分类

A5.96. 在有些情况下，分类方案由外部给出。例如，绘制贫困程度分区地图时，制图人员使用的是平均收入的特定临界值，即所谓的贫困线，低于该线的地区则是贫困地区。还有一个例子是，分类方案已经确定，要与一幅已经印制完毕的地图进行比较，但这些地图的原始数据得不到。例如，为了准确比较，比较一个国家各省的生育率地图，需要相同的分类方法。

#### 6. 一般性说明

A5.97. 从上文的概述中可以看出，有很多数据值归类方法。大多数地理信息系统和桌面地图绘制软件均支持等间距、分位点、标准偏差和自然中断方法。另外，所有软件都允许用户自己定义数据分类法。

A5.98. 各种方法都有优势和劣势，见表A5.1。究竟哪一种方法最合适，取决于数据分布情况和地图的目的。一般来说，数据分布应通过统计图表进行评估，例如上面提到的等级序列图。某些地理信息系统程序含有这些工具。最佳的类别数量以及最佳的中断点一般很容易找出。

A5.99. 应当指出的是，如果把几张地图放到一起比较(如一系列时间的地区性别比率、同一国家两个省份的安全用水地图)，自然中断法就不合适了。在这种情况下，各种类别之间的中断需要保持恒定。为此，应选用用户根据所有数据进行评估而自己定义的分类方案。如果目的仅是比较不同观测值随时间和空间而产生的升降，无需了解实际数值，则可使用分位点地图。例如，两幅四分法地图可以突出显示25%的地区最近一次和上一次人口普查的最高识字率。

表A5.1  
不同分类技术的评估

分类方法	优点	缺点
等间距法	容易使用。 适宜于均匀分布的数据。	在分类方案与数据分布之间没有关联。 因为类别间隔是固定的,相似的数值可能归入不同的类别,而不相似的数值可能归入同一类别。 不适合高反差的数据分布或具有界外值的数据集。
几何递增法	容易使用。 适宜于高反差的数据分布(如大量小数值和少量非常大的数值)。	必须由用户确定恰当的几何递增比。 因为类别间隔是固定的,相似的数值可能归入不同的类别,而不相似的数值可能归入同一类别。
分位点(等频)法	有良好的视觉对比。 适宜于相当均匀分布的数据。	类似的或同样的数值可能归入不同的类别。
标准偏差法	适于表现集中在平均值周围的发散趋势。 可将每个类别与整体平均值联系到一起。 适宜于正常分布的数据。	高反差分布或含有界外值的数据集(少量非常大的数值或非常小的数值)将导致使用大量的类别(如出现高于或低于平均值的几个标准偏差)。
自然中断法	相似的数值将归入同一类别。 类别的数目通常由中断点数目确定。	最终类别的范围可能非常不平均。 要求进行主观判断(依观察确定)。 无法对不同时期的地图进行比较。
无类别等值线图法	无需定义类别之间的中断点。 灰度或色调直接由数据值确定。 突出数据集内的连续数据分布。	大多数的输出设备只支持有限的可区别的灰度或色调。 具有不明显的灰度或色调差别的地图在复制时(如复印时)效果不好。 难以在大多数地理信息系统和地图绘制软件包里实现。

## D. 色彩选择

A5.100. 本《手册》讨论的地图实例都是使用灰度来表现符号。印制黑白出版物的成本低,并且在黑白复印机上复印时灰度地图容易辨认。但彩色让绘图人员有了更多的设计选择。彩色打印机和绘图仪的价格持续下降。另外,会有很多的地图出现在网站或电子出版物上。在这些地方的地图设计中,色彩将得到广泛应用,尽管仍应明智地使用。

A5.101. 为等值线地图定义彩色方案时,了解计算机解释色彩的原理十分有用。计算机定义色彩有几种模式。两种最常用的是色相-色调-饱和度(HVS)模式,以及红、绿、蓝(RGB)模式。“色相”一词指的是我们通常所说的颜色,例如“红”或“蓝”。从物理上讲,色相与反射光的频谱范围有关,可见光谱的范围从紫色开始,紫色的波长最短,然后是蓝、绿、黄、橙和红,红色的波长最长。“色调”有时也称为亮度(即色相-色调-饱和度(HLS))。色

调确定的是差别，例如淡粉与深红之间的差别，这两种颜色的色相是相同的。最后是“饱和度”，饱和度是亮度或强度的尺度。低饱和度的颜色显得黯淡，高饱和度的颜色看起来则很纯正。

A5.102. RGB模式以不同深浅的红、绿、蓝叠加组合到一起构成新的颜色。计算机屏幕和电视机屏幕均采用RGB方法。三种深度相同的颜色混合后形成灰度。最低级别的红、绿、蓝混合生成黑色，最高色调的红、绿、蓝混合生成白色。

A5.103. 颜色的选择取决于变量的测量级别、地图类型以及制图人员想要传达的信息。人类善于区分色相，不同色相使离散类别非常容易区别。例如，可以通过蓝色圆圈与红色圆圈对比，显示不同类型的学校。但在通过颜色区分地图符号时，必须考虑到色盲问题。色盲的人无法区别红和绿，这是最普遍型的色盲，涉及大约1%的男性，还有一种蓝黄色盲。还有一些人看不到光谱内的绿色。一般说来，不要在地图上使用红绿差别营造效果。

A5.104. 连续测量的变量，如人口、收入、比率和百分比等可用图形变量表示，构成特定的顺序。色调的差别(如从浅到深的同一种色相颜色)很容易让人联想到变量的变化程度。因此，较深的色调通常会让人联想到较高的数据值。例如，人口密度经常用红色调表示，低人口密度以浅红表示，高人口密度以深红表示。对于高反差的数据分布来说，色调并不与数据类别的数值直接构成比例关系。在上一节提到的人口密度变量例子中，如按照这种方法，很多非常浅的红色将被指定给大量的低数值，而非常深的颜色则指定给少数几个高数值。对于几何递增或类似递增类别则要使用步进式的色调差别加以表示。

A5.105. 如果一种分类方法包括了很多类别，制图人员完成的类别多得很难通过容易区别的色调进行打印。这时，可将相近的色相归并到一起，得出所谓的部分光谱颜色范围。这里再以人口密度为例，可以从淡黄开始，逐步过渡到橙黄乃至深红。这里需要考虑的一件重要的事情是，递增步进必须清楚，从较少的主色到较多的主色。如果地图将多种主亮色用于连续或顺序类别范围中的低值和高值，则不可能传达出明确的信息，导致观察者认识混乱。

A5.106. 还有一种应用是，不同色相可以适用于连续数据范围，这就是发散式数据等级。例如，一幅按行政单位显示的净移民数量地图。类别从大的负数(表示大量迁出)，然后通过零点，到达大的正数值(表示大量迁入)。为了突出大的负数和大的正数值表示的地区(移民对人口动态产生了很大的冲击)，可采用这样一种颜色方案，例如，从明亮的红色到浅红或粉红，然后经过白色(净人口迁移率为零)，最后用浅蓝和深蓝表示最高的人口迁入率。

A5.107. 最后谈谈如何绘制多变量地图的问题，即如何同时表现两个变量效果。例如，一副地图可以同时显示不同的识字率和生育率，所用的图例实际上是识字率和生育率类别可能得出的矩形组合。制图人员必须找到一种合适的色彩方案，例如，用部分光谱相邻的色相表示不同的识字率，用色调变化表示

生育率。遗憾的是，这样一幅地图很难理解。看图的人需要不停地查看图例，弄清色彩和两个变量的数据值之间的对应关系。因此，一般情况下，应该避免多变量地图。A5.6节将介绍另外一些表现多变量地理信息的方法。

A5.108. 现在回到我们在上文讨论的计量级的类型问题，表A5.2概括了有关灰度和色彩的一些使用原则(另见Brewer, 1994年)。

表A5.2

灰度级和颜色的选择

计量级	实例	黑白地图	彩色地图	
标称型	二值型	能否得到安全的饮用水(是/否)	黑和白, 或浅灰和深灰。	不同色相的强对比, 如蓝和红, 黄和绿。
	类别型	主要语言(英、法、西班牙语等)	具有类似视觉感受的不同图案。	具有相似色调和饱和度级别的不同色相, 没有任何先后顺序的暗示, 如蓝、绿、黄、紫等。
顺序型	学历(小学、初中等)	顺序的灰度, 各级间差别相对强烈。纹理不同, 能强调数据的顺序性更好。	某些色相或部分光谱范围, 类别间相差较大, 例如淡黄、橙色、中红、深红。	
离散型	家庭大小(1、2、3……人), 但不是平均的家庭大小	类似于顺序型数据, 但灰度级别间较小的差别可以接受。	类似于顺序型数据, 但灰度级别间较小的差别可以接受。	
连续型	序列型	识字率(可以是0%到100%之间的任何值)	范围连续的灰度差别。灰度级与数据值可以成比例, 也可以不成比例。灰度级别较小。但能够分辨清楚的差别可以接受。	在同一色相内或一部分光谱范围内连续的颜色范围。色调的微小变化可以接受。
	发散型	性别比率(低于1=女多于男; 高于1=男多于女)	必须运用纹理/图案的差别。一侧灰度填实, 另一侧使用纹理差别, 可达到较好的效果。	中和(白或灰)在中间, 范围连续的不同色相在两侧。如从浅到深的橙色用于低于1的数值, 而从浅到深的绿色用于高于1的数值。

## E. 地图图例的设计

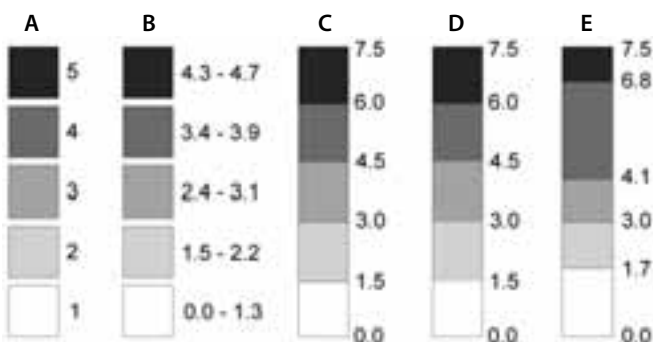
A5.109. 计量级可以反映在图例的设计中, 图例提供数据值或数据值范围和使用的图例符号间的基准。地理信息系统和桌面地图绘制软件有内置的图例设计功能, 能满足大多数应用的需要。然而, 对于更仔细的设计需要, 不论缺省图例在制图软件的版面模块内, 还是在外部图形软件中, 我们均可修改。

A5.110. 图A5.28列举了一些实例。对于类别数据, 每个图例方格应保持独立(见图A5.28A)。同样, 对于不连续的类型范围, 例如类型上限和类型下限之间有一个间隙时可用这种方式强调(见图A5.28B)。但在一般情况下, 应该避免使用这类图例。连续的图例方格强调变量的连续性, 例如比率或密度(见图A5.28C)。为了进一步强调数据的连续性, 每个类别方格不必用框线封闭起来(见图A5.28D)。最后, 带不规则类别范围的连续变量的类别图例见A5.28E。

A5.111. 图A5.28中的图例C、D和E表示的是中断点而不是独立的数据范围。使用数据范围代表连续性分布时会遇到这样的问题：同一个数据值出现在两个类别中：例如，0-10、10-20、20-30。可以通过使用“小于号”解决这个问题，使每一个值只出现在一个类别中，写成 $0 < 10$ 、 $10 < 20$ 、 $20 < 30$ 。对于一端开放的类别，可使用“大于或等于”符号，写成： $< 10$ 、 $10 < 20$ 、 $\geq 20$ 。

图A5.28

灰度地图内不同类型的图例



A5.112. 图例也可以与概括变量数据分布的统计图表集成在一起。把直方图中的条形涂成相应的灰度或颜色是常用的一种手段(见图A5.29A)。如果各类别的范围不完全相同，直方图中的条形可以画成宽度各异的条。如果地图绘制软件不支持直方图，可以在图形程序中先行设计好，或从电子表格软件或统计软件中制做出来后导入。有两种确定条形高度的方法。一个比较常规的方法是使用直属于各个类别的地理单元的数目。有些桌面制图软件把属于每个类别的单元数量显示在图例之内。这种方法有时会出现问题。例如地区的人口数差别极大。这时可放弃单元数量，用直方图的高度表示相关的人口数量。例如，对于一幅人口密度图来说，用直方图的高度表示生活在每一密度范围内的人口数量。当然，这样一来，直方图的形状会产生很大不同，需要在地图上清楚说明数据处理方法，或在配套的资料中加以说明。

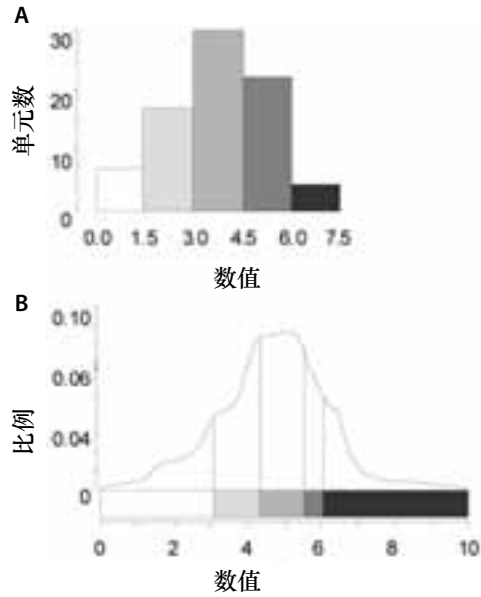
A5.113. 统计软件也可以通过密度计算图标表现比直方图更连续的数据分布情况(见图A5.29B)。在密度曲线下的表面积总和为一，因此每一单项数据值的大致频率可从图中看出。在《美国人口死亡率图册》中就采用过这种类型的图例(美国，国家健康统计中心，1997年)。

## F. 说明问题的地图

### 1. 多变量地图

A5.114. 除少数例外，前面的例子讲的都是每次显示一个变量的情况。这种方式在人口普查地图册中最普遍。有时出于分析和搞清楚变量之间关系的目的，也需要同时显示一个以上的变量。在有关颜色选择的部分曾经指出利用

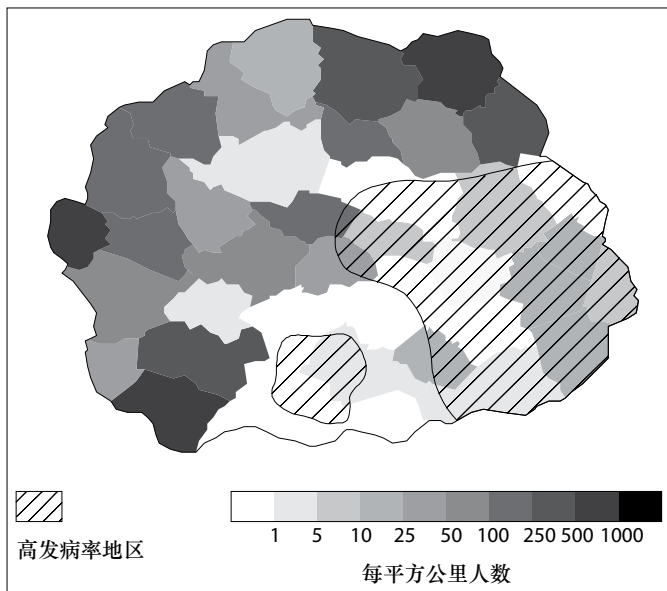
图A5.29  
用图例显示统计数据的分布



复杂的颜色方案在同一张地图上显示两个变量的多变量地图一般让人难以理解。还有一种选择在前面也提到过，就是将图案放到涂色的等值线图的透明背景之上。如果重叠变量只有少数几种，或干脆是二值变量(如是/否)，会取得很好的效果(见图A5.30)。

A5.115. 在统计数据分析中，仅就两个类别变量中的少量数值进行分析时可采用交叉式的表格。这种表格也称“关联表”。一个双向表格的行和列表

图A5.30  
将实色和填充线符号组合在同一地图上表现两个变量



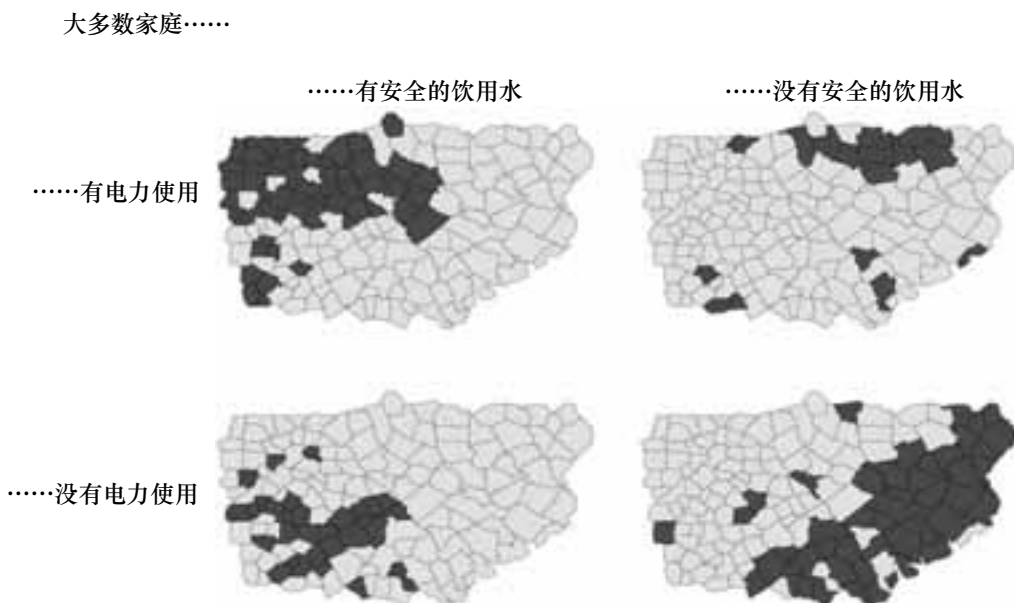
示两个变量的类别，单元格内显示取自每个变量相应值的观测值。这种安排方式便于快速判断变量之间的关系。例如，可以把住房普查得出的两个变量转换成两个二值变量，用以说明该地区大多数家庭是否能享受到公共设施服务。这两个变量一个是有安全饮用水的家庭百分比，另一个是有电力可使用的家庭的百分比。交叉表大致如下：

大多数家庭……			
	……有安全的饮用水	……没有安全的饮用水	总 计
……有电力使用	55	17	72
……没有电力使用	31	48	79
总 计	86	65	151

A5.116. 如果要表达这种地理信息，可以制作一幅具有四个类别的地图：每个类别代表交叉表内的一个单元格，但四个类别之间没有自然的先后顺序，观察者很难发现地图图案的差异。一种更好的方法是把双向表格直接转换成地图式表达手段。图A5.31显示了一幅与双向表格等价的地图。每一幅地图显示的地区相当于双向表中的一个单元格。这种地图不要求全面的图例，因为深色灰度已经能够清楚地强调所需表达的区域。

A5.117. 即使在一张很小的、大约三分之一页面的地图上，图案也能取得十分明显的效果。西北部的绝大多数地区均有安全饮用水和电力使用，而东南地区的大多数家庭则没有。在交叉表格中，非对角线的单元格通常是最有意义的。在某些东北地区，大多数家庭没有安全饮用水，但有电力使用。而西南地区的情况恰恰相反。

图A5.31  
与双向表格等价的地图



A5.118. 这种方法还可以扩大到更为复杂的表格，例如，一个变量取三个值(如低、中、高)，而另一个变量只包含两个类别。这种图没有必要画得很大。即使有很多地理单元(本例为151个地区)，小图也足够说明问题，因为只需要两种颜色或灰度。

## 2. 小型多幅地图

A5.119. 用多幅地图编排数据也可有效地表现动态信息。图A5.32根据四次连续人口普查图形说明人口的增长情况。人口密度图说明何处人口增长得最快。为了便于跨时间进行比较，类别的界限在所有的图上必须相同。也就是说，不宜采用基于数据分布的分类方案(如自然中断法)。人口密度图配有三幅小图，说明人口普查间隔期间的年人口平均增长率。

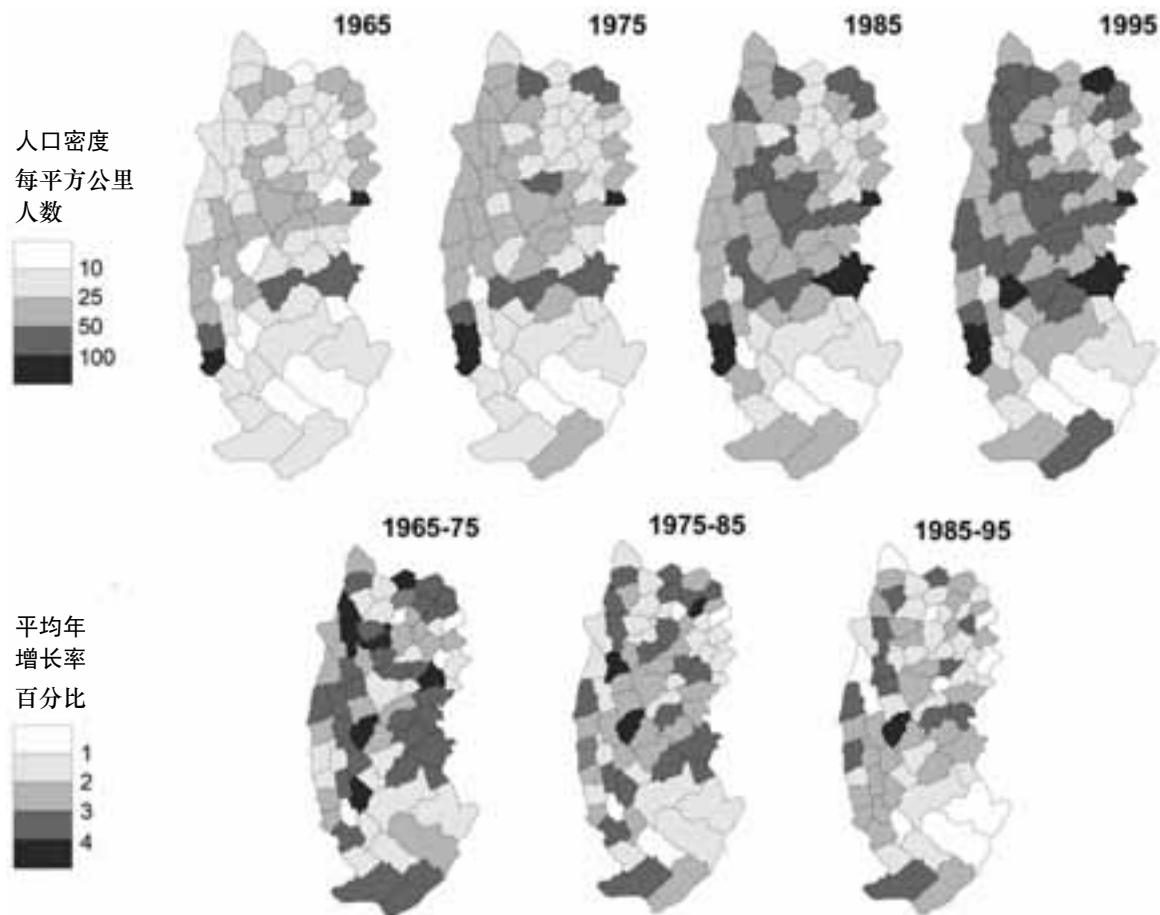
A5.120. 图A5.32中显示的那些地图，术语上叫做“小型多幅地图”(Bertin, 1973年; Tufte, 1983年)。同样的地图设计可反复使用，用于表示每一年或每一亚人口群发生的变化。由于所有的地图设计完全相同，观察者很容易理解其中的含义。因此地图设计者可以通过这种方法表示原本无法表示的更高密度的信息。通过小型多幅地图表达多个变量之间的关系常常比有着复杂的图例设计的复合地图要清楚得多。

A5.121. 另一个使用小型多幅地图概念的例子见联合国出版物《地理信息系统用于人口统计》(联合国, 1997年; 见图4.8)。其中以五年为一个年龄组说明尼泊尔境内75个地区的性别比率。整套图形包括17个小图，各种类别以性别比例平衡点为中心。在这幅地图的彩色版中，不同深浅的红色表示女性的冗余数，蓝色代表男性冗余数。黑白版中用实色表示女性的冗余数，用不同密度的点表示男性的冗余数。显然，颜色增强了地图表达信息的效果。虽然信息量很大，地图仍很容易读懂，因为类似数值集中在一起十分醒目。很明显，如果想从一张包含1 275(17x75)项数值的表格中获取同样的信息，和地图相比，使用表格要困难得多。事实上，这幅尼泊尔地图所显示的一些明确趋势，正是一些地区男人在生命周期中进行移民活动的结果。



图A5.32  
小型多幅地图——说明随时间产生的变化

1965-1995年人口动态



## 附 件 六

### 专业术语

**准确度 (Accuracy)**——测量结果或表达数值与实际、客观存在的数值的吻合程度。在地理信息系统项目中，首先要做的一部分工作是要确定可接受的准确度的具体要求，并且制定出准确度的标准。不可将准确度和精确度混为一谈。后者系指在测量工作中区分微小数量的能力。例如，一个点的位置可以精确地测量到(例如小数点后五个有效位)，但也许无法准确地进行测量(例如，可能偏离其实际位置好几米)。

**地址 (Address)**——分配给某个住宅单元、商用建筑物或其他任何建筑的号码或类似的标志。地址主要用于邮递服务，但对于行政管理也很重要。例如，在公民登记系统和人口普查中就要使用地址。

**地址匹配 (Address matching)**——利用街道地址，将一般属性信息匹配到一个街道网络的地理位置上。例如，将一个表格上的地址登记号码匹配到一幅综合性数字街道图上可以生成一个地理信息系统点层，从而显示出每个住户的位置。地址匹配有时也称为“地理编码”。

**行政单位 (Administrative unit)**——行使管理和行政功能的地理区。行政单位通常是通过法律行为确定和建立的。

**航空摄影 (Aerial photography)**——从空中平台(通常为低空飞行的飞机)拍摄照片的技术。有时也称为垂直摄影或正射摄影。航空照片可用于高准确度的摄影测量绘图。

**航空测量 (Aerial survey)**——借用航空摄影或其他遥感技术进行的制图测量。

**美国信息交换标准编码 (American Standard Code for Information Interchange (ASCII))**——为便于在计算机之间和跨越操作系统进行字母数字数据和特殊字符交换而开发的计算机编码。每个字符都被赋予了一个字节编码，即0-255之间的一个值。

**注释 (Annotation)**——用来在地图上标定地理特征的文本。注释可以储存在地理信息系统内，并可以绘制在地图上进行显示和打印。与属性表中的文本信息不同，注释仅用于制图显示而不用进行分析。

**弧 (Arc)**——见“线”。

**弧秒 (Arc second)**——经度或纬度的一秒，或者一度的1/3 600。

**区域 (Area)**——地球表面有一定界限的二维空间，它在地理信息系统中表现为一个多边形。

**区域插值 (Areal interpolation)**——将一个属性特征从一组报告区向另一组不兼容的报告区的转移。例如，根据区域人口的地理信息系统数据集对生态区内人口总数进行的估计即属于此种转移。

**区域单位 (Areal unit)**——一个常用于编辑和报告组合数据的自然或人工区域。例如：土地覆盖区或点查区 (EAs) 即为“区域单位”。

**属性 (Attribute)**——地理特征的特点。例如，可以连接到地理信息系统地理目标上的关系型数据库表格中所储存的数字信息域或文本信息域即为“属性”。一个点查区的属性可以是该区独特的标志符、以平方公里为单位的面积、总人口数或者住户数量等。地理属性与一般属性有时是有差别的。前者储存在与地理坐标文件紧密相连并含有多个信息域(如内部标志符、要素编码和面积等)的数据表格中。而一般属性则通常储存在可以与地理属性表格相连的单独数据表格中。

**自动制图/设备管理 (Automated Mapping/Facilities Management (AM/FM))**——地理信息系统在主要解决工程和维修问题的公共事业和公共工程领域的一种应用。

**频带 (Band)**——多光谱遥感图像的一个分层，它可以显示在一定电磁波谱范围内测得的信号。参见**多光谱图像**。

**带宽 (Bandwidth)**——可以通过通信连接进行转换的数字数据的数量。

**基本数据 (Base data)**——见**框架数据**。

**底图 (Base map)**——显示可用做位置参考的基本地理特征的地图。样本特征包括道路、行政区边界和住宅区。底图可用来编辑新的地理资料，也可以作为显示专题地图信息时的参考。

**基站 (Base station)**——全球定位系统的数据接收站，其位置已被准确、精确地测定，它可以为移动式全球定位系统接收器播送和(或)收集差分校正信息。参见**差分全球定位系统 (DGPS)**。

**北斗 (Beidou)**——中国提出的一个可替代全球定位系统的系统，将包括35颗卫星和服务，准确度最高为10米。

**二进制 (Binary)**——由两个数字组成或表示两个数字的数学进位制，如二元变量(例如“是”与“否”)中的一样。二进制同时也是计算机的一种编码形式，这种编码的基础是只采用两个数值(0和1)且被称为“二进制数位”的具体信息。

**比特、位 (Bit)**——表现为0或1数值的二进制数字。

**边界 (Boundary)**——能够限定面积单位的范围或者两个区域交汇位置的一条线。边界在地理信息系统中表现为一条线，这条线可以定义为多边形的一条

边。在地面上，边界可以是看得见的，也可以是看不见的。也就是说，边界可以是实实在在的特征，如道路和河流，也可以仅仅由地理坐标确定。

**每秒位数 (Bits per second (BPS))**——数字通信网络中表示传输速度的一种量度。

**缓冲区 (Buffer)**——地理特征(点、线或多边形)周围一定距离内的区或带。缓冲区运算是地理空间的基本性能之一。

**字节 (Byte)**——一个由8个2进制数字或数位构成、可通过计算机程序进行处理的单元组。千(K)字节大体由一千个字节组成，兆(M)字节由一百万个字节组成，吉(G)字节由十亿个字节组成。

**地籍信息 (Cadastral information)**——为司法和税务目的而对土地所有权方面的过去、现在和将来的权利与利益进行描述的记录。地籍图可以显示土地区块的地理位置和范围。现在许多国家在进行地籍测量时都使用地理信息系统来存储这种信息。地籍信息也称为“土地命名信息”。

**笛卡尔坐标系 (Cartesian coordinate system)**——二维空间呈直角相交的直线坐标系。笛卡尔坐标系可以提供—个框架，将位置以x/y坐标的形式精确定位。

**统计地图 (Cartogram)**——根据报告单位内记录的变量值将这些单位换算后绘制出的地图。统计地图也称为“变量值与面积关系图”。

**制图泛化 (Cartographic generalization)**——通过减少地图体现细节抽取实际世界特征的过程。这一过程包括挑选、分类、简化和符号化。

**制图法 (Cartography)**——对地球表面某一部分进行二维表达的艺术和科学。被表达的特征可能是真实的目标(地形图)，也可能是一些概念和更抽象的特征(专题地图)。

**人口普查地理框架 (Census geographic framework)**——人口普查机构在进行人口普查计数和数据制表时所用的地理信息收集和报告单位。该框架包括人口普查和行政单位的分级结构、表示符号和编码，以及不同单位之间的关系。

**中央子午线 (Central meridian)**——制图投影中定义x坐标原点的经线。

**矩心 (Centroid)**——一个多边形的数学中心。对于形状不规则的多边形而言，矩心可以看做是“重心”。

**链 (Chain)**——见线。

**通道 (Channel)**——捕获卫星信号的全球定位系统接收器中的电子组成部分之一。多通道接收器可以捕获和处理几个卫星同时发射的信号。

**导航图 (Chart)**——主要用于海上和空中导航用的图，如航海图和飞行导航图。

**等值线图 (Choropleth map)**——将报告单位中记录的数值分配给许多分散类别或类型的统计图。可利用为每个类别选用的符号(颜色或图案)将这些报告单位隐蔽起来。

**分类 (Classification)**——将目标划分成具有相同或相似特征的组的过程。在制图学上系指将符号分配给相似或具有相似值的地图特征。分类是用来简化地图以改善制图员的信息沟通。

**交换站 (Clearinghouse)**——从国家空间数据基础结构的意义上说，交换站是汇集和传播地理信息系统数据和元数据的仓库。

**客户端 (Client)**——使用存储在另一台(通常为远程的)计算机或服务器上的数据或软件的计算机称为客户端。

**代码 (Code)**——用于标记地理目标的字母数字符号。代码还可用来标记属性类型如人口密度范围、土地利用等级或各种产业等。参见**地理代码**。

**颜色模型 (Colour model)**——在计算机内用数字表达颜色的方法。例如，在RGB颜色模型中，颜色是用红、绿、蓝的数值级别来表示的。例如，纯红被定义为255, 0, 0。其他颜色模型还有HLS(色度、亮度、饱和度)以及CMY(青、品红、黄)模型。

**分色 (Colour separation)**——针对四种颜色(青、品红、黄、黑)中的每一种将图形文件分解成独立页面或文件的过程。分色是大多数专业印刷过程的基础。

**列 (Column)**——在地理信息系统的网络或光栅数据库中纵向排列的一组单元或像素。在数据库管理系统中，“列”是属性表格中的一个域或条目。

**计算机图形元文件 (Computer Graphics Metafile(CGM))**——用于图像或向量数据交换的一种标准文件格式。

**计算机辅助设计/计算机辅助设计和制图 (Computer-Aided Design/Computer-Aided Design and Drafting(CAD/CADD))**——为制图和设计，特别是为工程学或建筑学领域提供工具的软件系统。CAD系统采用图解坐标系，因此与地理信息系统类似。

**正形投影 (Conformal projection)**——一种在每一个点上都正确保存了所有角度的制图投影。

**连接性 (Connectivity)**——在拓扑式的地理信息系统中，两条或多条曲线在一个单一点或节点上连接的性质。

**邻接 (Contiguity)**——两个或多个地理特征彼此相邻或毗邻。

**连续地理现象 (Continuous geographical phenomena)**——在发生变化时，无明显间断或中断的地理变量，如温度和大气压。“连续地理现象”与“离散地理现象”是相对的概念。

**等高线 (Contour)**——在地图上将高度相等的点连接起来的线。参见**等值线**。

**控制 (Control)**——见**大地测量控制**。

**控制点 (Control point)**——在地图或航摄照片上,或在数字数据库中,已知x, y 坐标(可能还有高度)的一个点。控制点用来对地图特征进行地理学方面的登记。

**控制部分 (Control segment)**——系指全球定位系统监测网络和控制站组成的一个全球网络,以确保卫星信号的准确性。

**坐标 (Coordinate)**——在二维或三维系统(例如x/y或者x/y/z,其中z代表高度)中描述一个点的位置的两个或三个数值。二维坐标有时称为“成对坐标”,三维坐标有时称为“三联坐标”。在地理信息系统数据库中,坐标代表地球表面上相对于其他地点的位置。

**坐标几何学 (Coordinate Geometry(COGO))**——土地测量员在对位置进行精确测量中所用的术语。

**坐标系 (Coordinate system)**——在地图上或地理信息系统数据库中用来标定位置的基准系统。制图坐标系通过地图投影、基准椭圆、中央子午线、一条或多条标准纬线、可能还有x和y的坐标值更改等确定。

**覆盖层 (Coverage)**——在地理信息系统中,覆盖层有时指一个向量数据集,在这个数据集中包含属于单一专题(如人口普查单位或道路)的地理特征。

**数据捕获 (Data capture)**——将硬拷贝源中或者借助现场测量的地理坐标数据转换成为计算机可读取格式的过程。数据捕获通常包括纸质地图或者航摄照片的数字化处理或扫描。

**数据转换 (Data conversion)**——将数据从一种格式转换成另一种格式的过程。通常情况下,数据转换系指将纸质地图信息转换成数字形式。从广义上说,地理数据转换还包括将数字信息从一种地理信息系统文件格式到另一种地理信息系统文件格式的转换。

**数据字典 (Data dictionary)**——描述数据库内容的数据目录。信息是围绕着属性表格中的每一个信息域以及属性表格的格式、定义和结构编排的。数据字典是元数据信息的一个重要组成部分。

**数据格式 (Data format)**——通常指软件系统内特定的、可能获有专利的一套数据结构。

**数据模型 (Data model)**——系指用户对一个可以描述数据库各个实体及彼此关系的数据集所进行的概念化设计。

**数据集 (Data sets)**——系指涉及单一专题的数值或数据库的目标逻辑加总。

**数据标准化 (Data standardization)**——在共用的数据定义、格式以及各种数据层和数据要素的表达和结构等问题上达成一致的过程。

**数据结构 (Data structure)**——数据模型的执行。这种数据模型由许多用来代表各种地理特征的文件结构组成。

**数据类型 (Data type)**——属性表格中列的信息域特征，例如字符、浮点以及整数等。

**数据库 (Database)**——系指彼此相关并作为一个单元管理和存储的信息逻辑加总(例如在同一个计算机文件中)。数据库和数据集这两个术语常常可以互换使用。地理信息系统数据库含有实际地理特征的位置及其特点的信息。

**数据库管理系统 (Database Management System(DBMS))**——设计用来对表列数据进行管理和操作的软件包。该系统可用于数据的输入、存储、操作、检索和查询。多数地理信息系统都使用关系型数据库管理系统来管理属性数据。

**基准 (Datum)**——制图学上定义坐标系的一套参数。具体而言，基准是测量和计算的参照物或基础。例如，一个国家的制图基准可以为该国的制图工作提供参照框架。

**差分全球定位系统 (Differential GPS(DGPS))**——在一个已被准确和精确确定的位置上，通过计算第二个全球定位系统接收站(基站)的信号误差(偏移量)，以提高全球定位系统捕获坐标的准确度的一套技术。对于由移动式装置捕获的坐标可以采用校正系数进行校正，校正的方式既可是实时的，也可以采取后处理方式(即利用经过时间标定的校正信息数据库)。在世界上某些地区，差分校正信息是从一系列永久性基站连续播送出来的。

**数字高度模型 (Digital Elevation Model(DEM))**——地球表面某一部分的高度信息的数字表达。数字高度模型通常是一个光栅数据集，数据集中为精细网格的三元存储了高度值。但也可以用向量格式存储高度。数字高度模型有时也称为“数字地形模型”(DTM)。

**数字正射照片 (Digital orthophoto)**——通常是具有极高的分辨率并经过几何校正的数字图像或航摄照片。数字正射照片也称为正射影像，它兼顾了航摄照片的清晰度和地形图的几何准确度。

**数字地形模型 (Digital Terrain Model(DTM))**——见“数字高度模型”(DEM)。

**数码化表 (Digitizing table)**——用来从纸图或类似的制图资料中捕获坐标数据的计算机外围设备。数码化表也称为“数字化仪”。

**数字化处理 (Digitizing)**——将纸质地图上的地理特征信息转换成数字化坐标的过程。数字化处理一般指利用能够捕获坐标并将其存储在地理信息系统数据库中的鼠标状光标，在附于数码化表上的纸图上进行人工描线的过程。

**离散地理特征 (Discrete geographical features)**——容易区分开的单个实体，例如房屋或道路。离散地理特征和连续地理现象是两个相对的概念。

**分解(Dissolve)**——地理信息系统的功能之一，能将相邻的、对于特定属性来说具有相同值的多边形边界进行删除。例如，点查区多边形可以按其管理单位编码进行删除，从而产生监管图。

**点图(Dot map)**——把数量或者密度用点表示出来的地图。一般情况下，每一个点都代表着确定好的离散目标(如人或牲畜)数量。这些点可以随机布置在报告单位内，或用来说明变量基本的真实分布情况。

**图形交换格式(Drawing Exchange Format(DXF))**——由加利福尼亚索沙利托的Autodesk公司开发的，用于描述图形或图案的ASCII格式。这种格式在开发之初是为了用于计算机辅助设计，现已变成地理信息系统数据交换的一种标准。

**边缘匹配(Edgematch)**——地理信息系统中的一种手工或自动化编辑技术，可将相邻图幅中数字化的共用地理特征匹配在一起。边缘匹配是一种必要的技术，例如，当几幅分别进行了数字化处理的地图进行拼合后，就必须采用这种技术将道路或者行政单元的边界连接起来。

**椭圆球体(Ellipsoid)**——在制图学上用以代表地球的三维形状。地球椭圆球体从中心到两极的距离(半短轴)要小于从中心到赤道的距离(半长轴)。椭圆球体也称为“旋转椭圆球体”。

**实体(Entity)**——某种类型的真实现象。在数据库管理系统中，系指共享相同属性的目标(例如人或住所)的加总。实体是在数据库初步设计阶段确定的。

**实体关系模型(Entity-relationship model)**——定义实体及其彼此关系的数据模型(例如，定义点查区与监控区之间的关系)。

**点查区(Enumeration area(EA))**——用以汇总、编辑和发布人口普查信息的，通常为最小的地理单位。点查区可以用草图上或者地理信息系统数据库中所描述的边界确定。这些边界在地面上可以是可见的，也可以是不可见的。点查区也称为“人口普查区块”或“人口普查区”。

**等面积投影(Equal area projection)**——一种制图学投影，投影时所有的区均按照与实际地区的正确比例进行显示。

**赤道(Equator)**——制图学上的一条基准纬线，即南北纬度均为“0”的纬线。

**等距离投影(Equidistant projection)**——沿着图上一条或多条曲线或从图上一两个点到所有其他点始终保持一定比例的一种制图学投影。

**地理特征(Feature)**——图上显示的或在地理信息系统数据库中存储的地理目标。地理特征可以是自然特征，也可以是人造的真实目标(一条河或一个住宅区)；同时也可以是想象的或定义的特征(如行政区边界)。

**域(Field)**——数据库表格中的列。

**文件传输协议(File Transfer Protocol(FTP))**——数字通信系统(如因特网)中交换计算机数据文件的一套标准。



**流动图 (Flow map)**——显示货物或人员等沿着一条线性轨迹运动的地图。

**外关键字 (Foreign key)**——在关系型数据库管理系统内，一个表格中所含的可识别另一表格“行”值的域或条目称为外关键字。通过确定关系型数据库中两个元素之间的关系，可以用外关键字将两个表格连接起来。外关键字是另一表格中的主关键字。

**框架数据 (Framework data)**——系指在国家地理信息系统活动中所用的一套通用用途的地理专题或基本数据，例如行政区的边界、高度或者交通基础设施等。提出框架或国家空间数据基础结构这一概念的目的是为了协调一个国家框架数据的地理信息系统数据集的开发和标准化工作。

**伽利略 (Galileo)**——将由欧洲联盟建造的全球定位系统的替代系统。该系统将包括30颗卫星和两个地面站，并且在用户层面可与美国全球定位系统兼容。

**地名字典 (Gazetteer)**——地名及地理位置(通常为经纬度)清单。

**泛化 (Generalization)**——见制图泛化。

**地理编码 (Geocoding)**——(a) 在数字数据库中将地理代码分配给地理特征的过程；(b) 一种根据街道地址确定一个点位置的地理信息系统功能，参见地址匹配。

**大地测量控制网 (Geodetic control)**——由经过精确和准确测量的控制标志或参照标志组成的一个网络。这些标志可作为新位置测量的基础。大地测量控制网也称为“基准点”。

**地理属性文件 (Geographic attributes file)**——与存储在地理信息系统坐标文件中的空间目标紧密相连的一个数据库表格。地理属性文件或表格含有每个地理特征的特定信息，例如该地理特征的标志符、名称和表面积。某些系统中，这种文件也称为点、线或多边形属性表。存储在外表中的数据可以通过关系型的数据库操作连接起来。

**地理代码 (Geographic code)**——指定给一个法定单位、行政单位、统计单位或者报告单位的一种独特的字母数字标志符。

**地理数据库 (Geographic database)**——与地球表面位置有关的地理特征数据的逻辑加总。

**地理谱系 (Geographic hierarchy)**——在人口普查地图绘制过程中，将用于行政管理或数据采集的面积单元按通行方式套叠在一起的系统。例如，一个国家可分为若干个省，省又可再分为若干个区，以此类推直至最低的级别，这个最低级别可能就是点查区。参见人口普查地理框架。

**地理信息系统 (Geographic information system(GIS))**——将计算机硬件、软件、地理数据和人员组合在一起的一套系统，用于获取、存储、检索、更新、处理、分析以及显示地理参考信息。

**地理目标 (Geographic object)**——由用户确定的，可以在地理数据库中进行表达的地理特征或地理现象，例如街道、土地区块、水井、湖泊等。

**地理基准文件 (Geographic reference file)**——一种以表格形式存在的数字主文件，列有与人口普查和测量数据采集有关的所有地理实体的名称和地理代码，可能还包括其属性。

**地理相符 (Geographically coincident)**——两个或两个以上地理特征共用同一个位置或同一条边界的性质。例如，有些报告单位或统计单位同时也可能是行政管理单位。

**地理基准 (Georeferencing)**——确定页面坐标与实际坐标之间关系的过程。进行数字化处理后必须进行地理基准工作，例如将以数字化单位(如厘米或英寸)量度的页面坐标转化成用来绘制原始图件的实际坐标系。另见**转换**。

**地理空间的 (Geospatial)**——一个有时用来描述地理信息或空间信息的术语。

**地理静态卫星 (Geostationary satellite)**——在地球表面一个点的上方保持固定位置的地球卫星，也称为“地球同步轨道卫星”。

**地理标记图象文件格式 (Geo-TIFF)**——见**标记图像文件格式**。

**全球导航卫星系统 (Global Navigation Satellite System(GLONASS))**——由俄罗斯联邦国防部管理的，与美国的全球定位系统相当的导航卫星系统。该系统与全球定位系统非常相似，但两者无法换用。有些接收装置将全球定位系统和此系统的信号结合起来使用以提高坐标准确度。

**全球定位系统 (Global Positioning System(GPS))**——由绕地球运行的，发射用来准确测定地球表面地理信号的24颗卫星组成的系统。全球定位系统广泛用于野外测绘、测量和导航。全球定位系统由美国国防部管理。参见**差分全球定位系统**、**北斗**、**伽利略**和**全球导航卫星系统**。

**政府单位 (Governmental unit)**——见**管理单位**。

**分级符号 (Graduated symbols)**——在专题制图学上，用来代表一个点或者一个报告单位内变量级的符号(如圆圈或正方形)。符号的大小与变量值成正比。

**可交换图像文件 (Graphic Interchange File(GIF))**——一种图形图像的文件格式，最初的开发目的是要借助电子公告牌传输图像。利用这种文件格式可以有效地压缩文件尺寸。目前网页上大多数图形图像均采用这种格式。

**格线 (Graticule)**——系指制图学中绘制在图上的，由经纬线组成的网格。

**大圆 (Great circle)**——球被穿过球体中心的平面截成的圆称为大圆。例如，所有的子午线和赤道均是大圆。球面上两点间的最短轨迹是沿着穿过这两点的大圆上的轨迹。

**格林尼治子午线 (Greenwich meridian)**——作为基准的经线，即东经和西经均为“0”的经线，该经线穿过伦敦郊区的格林尼治镇。

**网格 (Grid)**——一种地理数据模型，代表以均匀方格排列成网状的信息。网格上的每个方格都有一个数值，这个数值代表该位置上的一种地理现象(如人口密度或者温度)的实际值；网格的每个方格也可以表示一种类别或种类(如点查区标志符或土壤类型)。参见**光栅**。

**地面真值 (Ground truth)**——野外测量中采集的信息，用于验证或校准遥感数据中抽取的信息。

**平视数字化处理 (Heads-up digitizing)**——不使用数码化表的一种数字化技术。它是利用屏幕鼠标，从背景显示的扫描图像上追索地理特征，或者是追索与计算机屏幕相连的透明介质(如聚酯薄膜)上所绘制的地理特征。

**水文图 (Hydrography)**——与地表水(如湖泊、河流和水渠)有关的地理特征。

**地势图 (Hypsography)**——与地形起伏或高度有关的地理特征。

**图像 (Image)**——地球表面某一部分的显示。但是，图像通常是利用光学或电子传感装置生成的。例如，扫描的航空照片或者遥感数据通常可称为图像。就数据存储和处理而言，图像与光栅或网格极为相似。

**基础设施 (Infrastructure)**——一个国家、州或地区的公共工程系统，包括道路、公共事业管线和公共建筑物等。

**集成 (Integration)**——系指地理信息系统中，将来自非均质源的空间数据汇编成一个协调一致的数据集的过程。垂直积分系指地理信息系统将不同的，但以同一个坐标系定位的数据层组合在一起的能力。

**因特网 (Internet)**——全球计算机联网系统。它可以提供各种数据通信服务，例如远程登录、文件传输、电子邮件、公告牌和新闻组等。因特网也是万维网的基础。

**因特网协议 (Internet Protocol(IP))**——一套最重要的编码和规定，它们使通过因特网进行数字数据传输成为可能。

**插值 (Interpolation)**——根据相邻位置上测得的值估算某位置上的变量值的过程。插值是用来根据点样信息求出一个完整的网格数据集(例如根据降雨测量站数据求出降雨面积)的过程。

**交叉 (Intersecting)**——地理信息系统的一种功能，用来从地形学角度将两个空间数据层加总或者综合在一起，以便只保存这两个数据层共有区内的地理特征。

**等值线 (Isoline)**——在所谓的等值线图上将数值相同的一些点连接起来的线。最明显的例子是等高线，即显示出相等的高度(也称为等高线图)。

**Java语言 (Java)**——一种计算机编程语言，用以生成可在多个平台(即操作系统)上运行的软件包。**Java程序**(即小程序)可以通过因特网在远程计算机上发送或检索。

**连接 (Join)**——在关系型数据库管理系统中，将一个外关键字链接到外部表格中该码的主例上之后，将数据库表格中的值附加到另一表格上的过程。

**联合图像专家组图像格式 (Joint Photographic Experts Group(JPEG))**——这是一种主要用于摄影图像的图形文件格式，可以大幅度压缩文件的大小。

**土地信息系统 (Land Information System(LIS))**——有时在地理信息系统应用中使用的术语。该系统含有某个特定地区的信息，其中包括地籍资料、土地利用情况以及土地覆盖情况等。

**纬度 (Latitude)**——球面上极坐标系的“y轴”。纬度以赤道南北两侧角度的度数进行量度。纬度线也称为纬线。

**层 (Layer)**——包含有属于同一专题地理特征(如道路或房屋)的单个地理信息系统数据集。该术语说明地理信息系统对处于同一坐标系的不同专题层进行叠加和合成的能力。另见**覆盖层**。

**图例 (Legend)**——在制图学上，用来解释图上所示地理特征和变量的符号信息。其中包括用来解释地图的符号图案，例如人口密度图上使用的浓淡不一的颜色和相应值的范围。

**线 (Line)**——一种一维目标。它是由一系列x, y坐标构成的地理数据类型，其中第一个和最后一个坐标称为节点，中间坐标称为极点，有时线也可指一条弧或一条链。一条线与其他线相交形成的两个交点之间的部分称为**线段或弧段**。

**多边形内线法 (Line-in-polygon)**——地理信息系统的一种运行方法，运用这一方法可以将线特征与多边形特征结合起来，以确定哪些线段落落在哪些多边形内。用这种方法还可将多边形的属性加到线表格中每一种相应的记录中(例如将道路放在一个区内)，同时也可求出每个相应多边形线属性的总数(例如一个区的道路总长度)。

**局域网 (Local Area Network(LAN))**——将计算机在一个较短距离内(如同一座办公大楼内)连接起来形成的计算机网络。

**逻辑准确度 (Logical accuracy)**——对图上或地理信息系统数据库内多个地理特征之间的关系正确表达(如相邻或相连)的程度。一个地理信息系统数据库，即使其位置准确度有限，其逻辑准确度也可能很高。

**经度 (Longitude)**——球面上极坐标系的“x轴”。经度是以格林尼治子午线东西两侧的角度度数进行量度。

**地图 (Map)**——通过在一个平坦表面(如纸张或计算机屏幕)对地球表面的某一部分进行描绘的表达方式。

**地图编制 (Map compilation)**——将制图测绘结果与资料进行组合、评估和解译以形成一幅新图的过程。

**地图合成 (Map composition)**——将地图元件装配在一起，以形成一件制图产品的过程。该产品应当样式美观且能正确显示所代表的地理现象。

**地图元件 (Map elements)**——专题地图或地形图的组成部分，如标题、图例、比例尺、指北箭头、网格线、边界线、图幅框边线等。

**图幅范围 (Map extent)**——图解单元中用来确定一个矩形的坐标。这个矩形包含了一个特定图解显示器中或地理信息系统数据库中所含的一切要素。也就是数字数据库中的最小和最大x和y坐标，或者图解显示器上显示的数据库的一部分。

**地图投影 (Map projection)**——将地面位置换算成平面坐标系的一种数学过程。根据采用的数学公式的不同，地图投影的性质也有所不同。有的投影保留了地球上各个地区的形状，有的则保留相对面积、角度或者距离。

**地图单位 (Map units)**——在地理信息系统数据库中存储的坐标的测量单位，例如厘米、米或者度、分、秒等。

**子午线 (Meridian)**——由相应的经度确定的参考线，例如格林尼治子午线。

**元数据 (Metadata)**——关于数据的数据。可以描述一个数据集的内容、质量、条件、格式、谱系及其他相关特征信息的加总。

**最小制图单元 (Minimum mapping unit)**——一般说来，系指将要绘制在地图上的最小地理特征的大小。在给定地图比例尺时，也可表示把细小的紧凑的多边形特征显示为一个点或狭长多边形特征显示为一条线那样的尺寸或大小。例如，如果一个城市在页面上的尺寸大于3毫米，则可显示为一个多边形；如果小于3毫米，则可显示成一个点。

**多路径误差 (Multipath)**——全球定位系统信号在相邻构筑物(如房屋或树木)上反射和散射引起的全球定位系统读数误差。多路径误差一般出现在高精度测量中。

**多光谱图像 (Multispectral image)**——由许多频带或分层构成的遥感数据集。这些图像是在同一时间为同一地区拍摄的，并且基本上是各自独立的图像，每个图像都显示出电磁谱的一个不同的谱段。

**天底 (Nadir)**——在航空摄影和遥感测量中，直接位于摄影机镜头或者传感器下的地球表面的一个点。

**网络分析 (Network analysis)**——对地理信息系统数据库中代表某种目标(如街道网络)的一组曲线上的各个点或地址之间关系进行分析的方法。网络分析可用以确定位置并在紧急管理等情况下选定路线。

**节点 (Node)**——一条线性特征的起始点或终结点，或者两条或多条曲线的交点。

**正常化 (Normalization)**——数据库设计中的一种构思方法。利用正常化方法，通常建立数据库多个实体间的彼此依赖性和关系，可以剔除一个复杂数据库的冗余信息。正常化可减少存储量，避免数据库中的不和谐。

**正射照片 (Orthophoto)**——见**数字正射照片**。

**叠加 (Overlay)**——处于同一个地理参照系中的两个数据层的合并。可以为制图学显示目的进行叠加，也可以将两个数据层进行真正的合并而创建一个新的地理信息系统数据集(例如多边形叠加，或者多边形内点和多边形内线的合并)。

**出头 (Overshoot)**——在数字化处理过程中，系指将一条线延长到了它本应与另一条线相交的交点之外。由此产生的虚伪线段有时也称为“悬空线”。

**全色图像 (Panchromatic image)**——在很宽的电磁谱范围记录下信号的一种遥感图像。全色图像类似于黑白照片。

**区块 (Parcel)**——单一的地籍单位或地产。

**摄影测量学 (Photogrammetry)**——从照片中抽取测量信息和其他信息的艺术和科学。在绘图过程中系指从航摄照片或卫星图像中抽取实际地理特征信息的方法。

**像素 (Pixel)**——即照片要素，类似于图像、网格或光栅图像中的一个单元。

**平面坐标系 (Planar coordinate system)**——确定位置的一种坐标系，该系统中两组直线呈直角相交且以一个选定的垂直交点作为其原点。参见**笛卡尔坐标系**。

**平面图 (Planimetric map)**——与地形图不同的，仅显示地理特征的位置而非其高度的一种平面图。平面图可以显示与地形图所示相同的地理特征(地形和高度等值线除外)，但通常只显示为特定目的而选用的某些特征。

**绘图仪 (Plotter)**——可以绘制图形文件的计算机外围设备，类似于打印机，但一般输出版面更大。

**点 (Point)**——尺寸为零的目标。数字地理数据库中用来代表那些尺寸太小无法以线或多边形形状显示的地理特征的x, y坐标。例如，住宅、水井或大楼常常以点显示出来。

**多边形内点法 (Point-in-polygon)**——这是一种地理信息系统运算方法，即将点要素与多边形结合使用以确定哪些点位于哪些多边形内。运用这种方法，可以将多边形属性加到点属性表格中每一种相应的记录上(例如将卫生服务区信息加到测量抽样点上)，也可以求出每个相应多边形的点属性总数(例如每个地区的医院总数)。

**多边形 (Polygon)**——一种二维目标，即在向量地理信息系统中表现为一个序列系列的x/y坐标的面积特征。这些坐标点确定包围这一面积的边线，也就是说多边形的第一个和最后一个坐标是相同的。

**多边形叠加 (Polygon overlay)**——一种地理信息系统运算方法，即将两个多边形数据层结合在一起，创造出一个新的数据层的方法。输出层由两组输入多边形的叠加区构成。这个新数据层的属性表格包含了这两个输入数据集的属性。多边形叠加是地理信息系统的基本运算方法之一，常用来将从非均质源得来的信息(如人口统计数据和环境数据)进行综合。

**位置准确度 (Positional accuracy)**——系指地图上或地理信息系统数据库中的位置相对于其在地球表面的实际位置正确记录的程度。逻辑准确度则不同，它指的只是表达地理特征之间关系的正确程度。

**外部页面语言 (Postscript)**——这是一种灵活的、高分辨率的页面描述语言，多用来将图形信息(如地理信息系统生成的地图)传送给打印机。压缩式外部页面格式 (EPS) 包括为预览目的使用的小位图显示方式。

**精确度 (Precision)**——测量中区分微小差异的能力。在地理信息系统中，坐标的精确度取决于用来存储x和y坐标的数据类型(通常为双倍精度，或者每个数字用16个字节)。

**主关键字 (Primary key)**——属性表格中专门用来区别特定实例、行或记录的一个或多个域。

**协议 (Protocol)**——电子通信系统中决定如何进行数据处理、数据交换和数据格式化的一套规则。类似于一种数据标准，但只适用于操作过程。

**四边形 (Quadrangle)**——由成对的经线和纬线限制边界的矩形区。

**质量控制 (Quality control)**——数据库开发项目或制图生产系统中为了保证所产生的数据或输出能够符合规定的精确度标准和适用标准而采取的步骤和方法。

**分位点 (Quantile)**——一种统计学或制图学分类方法，即将同等数量的目标划分成固定数量的类别。四类别系统称为四分位点，五类别系统称为五分位点，十类别系统称为百分位点。例如，数据分布的四个四分位点中的第一个含有最低观测值的25%。

**半径 (Radius)**——从一个圆的中心至其外边缘的距离。

**光栅 (Raster)**——代表呈规则行列排布的信息的一种地理数据模型，类似于网格或图像。光栅的格子通常是，但并非都是正方形。面积特征或线条特征可以用相同值的相邻光栅格的组合来表达。

**调整 (Rectification)**——将图像或网格从图像坐标向实际坐标转换的过程。调整时通常要对网格单元进行循环计算和换算，因此需要重新取样或进行网格值的插值。它类似于向量数据的**转换**。

**对准 (Registration)**——将两幅图或两个地理信息系统数据层的特征进行匹配使之与对应目标吻合的过程。对准以一系列地面控制点为基础且涉及**转换**和**橡胶板法 (rubber-sheeting)** 处理。

**基准地图 (Reference map)**——在人口普查地图绘制过程中，显示人口普查地理框架某一部分的地图产品(硬拷贝或数字化地图)，例如一个数据加总单位或者一个统计发布单位。

**关系型数据库管理系统 (Relational Database Management System(RDBMS))**——根据一个共有域(主关键字和外部关键字)可以将几个数据表格临时或永久连接起来的数据库管理系统。数据库中每一行、每项纪录或每个实例都有一套固定的属性或域。每个表格都有一个专门识别每项记录的主关键字。这个表格也可能会有一个与外表中的主关键字相同的外部关键字。通过将外部关键字的值与外表主关键字中的相应值进行匹配实现关系连接。

**遥感 (Remote sensing)**——从一定距离以外(即无物理性接触)获取某目标信息的过程。遥感通常指利用卫星传感器或航空摄影技术获取图像的过程。

**分辨率 (Resolution)**——系指图形或数字数据库能够区分的最小细部的量度。在给定地图比例尺的情况下，分辨率决定了地图上地理特征的位置和形状的准确表示程度。对于光栅地理信息系统和图像数据而言，分辨率有时用来表示光栅大小或像素的大小。

**行 (Row)**——在地理信息系统中，系指一个网格或光栅地理信息系统数据库中呈水平排列的一组单元格或像素。在数据库管理系统中则指属性表格中的一项记录或者一个实例。

**橡胶板法 (Rubber-sheeting)**——一种通过非均匀方式修改地理信息系统数据库中目标形态和位置的方法。橡胶板法常用来将未知坐标系中的地理信息系统数据集置于已知的坐标系中。通过从输入数据集的位置到输出坐标系相应正确的基准点或控制点建立一系列链接实现这种调整。

**运行长度编码 (Run-length encoding)**——光栅、网格或图像数据的压缩技术。该系统不是存储具同一值的相邻单元的每一个值，而是存储那个相同值及该值重复的次数。当光栅地理信息系统中存储离散目标时，这种压缩技术十分有用。

**卫星图像 (Satellite image)**——通过卫星上安装的摄影机或扫描器从地球轨道卫星上记录到的数字数据集。地理信息系统的卫星图像类似于光栅或网格数据集。

**比例尺 (Scale)**——在制图学上指图上距离与地面相应距离之间的关系。比例尺是以比值的形式实现的。例如1: 100 000意味着地图上的1厘米等于地球表面上的100 000厘米。既然比例尺是一种比值，那么小比例尺地图可以显示较大的地区，大比例尺地图则可显示较小的地区。从更普遍意义上说，比例尺表明了观测或调查的程度，其对象可以从微观规模的现象到宏观规模的现象不等。

**扫描 (Scanning)**——一种数据捕获技术，即借助光敏光学装置捕捉硬拷贝文件(如纸或聚酯薄膜文件)上的信息并将其转换成数字图像的过程。对于图



形数据来说，扫描是数字化数据输入的一种替代办法。对一幅地图进行扫描后，利用光栅-向量转换软件或者在屏线点特征跟踪技术，通常可以将图像数据转换成向量格式。

**示意图 (Schematic map)**——见草图。

**选择可用性 (Selective availability)**——美国国防部故意使用的，旨在降低全球定位系统卫星信号准确度的一种技术。该项技术已在2000年停用，但战时可恢复。

**服务器 (Server)**——为向其他计算机(客户端)提供某种服务而使用的计算机。例如，网络服务器是万维网数据、软件或内容的中央存储器。

**草图 (Sketch map)**——通常为手绘的，显示给定地区主要地理特征的地图。这种图的位置准确度不高，因此可能无法正确显示目标的距离和大小。但是草图具有较高的逻辑准确度，也就是说它能正确地显示目标之间的关系。草图也称为示意图或线条图。

**源资料 (Source material)**——用来编辑地图或地理信息系统数据库的任何类型的数据和信息。源资料包括与自然和人造地理特征有关的野外观测资料、空中和地面拍摄的照片、卫星图像、草图、专题地图、地形图、水文地图、等高线图、草绘地图、图表、表格信息、书面报表等。

**太空段 (Space segment)**——全球定位系统位于太空的那一部分，即24颗全球定位系统卫星。

**空间分析 (Spatial analysis)**——从以地理特征为基础获得的数据中抽取有用信息的一种技术。空间分析包括地理数据集集成、数据的定性和定量评价以及模拟、解译和预测。在地理信息系统中，空间分析常指地理信息系统数据集成的方法，如多边形叠加或者邻域分析。广义上说，它包括多种空间过程模拟(如迁移动力学模型)和空间统计学模型(例如用来解释观测值的空间排列和相互关系的回归模型)。

**空间数据 (Spatial data)**——有关地理特征的位置、大小、形态及相互关系的信息。在地理信息系统中，空间数据在技术上可以划分为点、线、面以及光栅网格。

**空间数据基础结构 (Spatial data infrastructure)**——参见框架数据。

**空间数据传输标准 (Spatial Data Transfer Standard(SDTS))**——在数据生产者 and 使用者之间以及软件系统和文件格式之间进行地理信息系统数据集交换的一种数据和元数据标准。现在已在执行或已经提出了许多国家标准和国际标准。

**空间指数 (Spatial index)**——地理信息系统或数据库管理系统为了加快空间要素的查询、分析运算和显示而采用的地理数据库中的一种供查找用的表格或配置。

**空间相互作用 (Spatial interaction)**——地理实体之间的相互依赖。空间相互作用常指货物、服务、信息或人员在各个地理位置之间的流动。空间相互作用分析是人员迁移研究中的一项重要工作。

**球体 (Sphere)**——类似于球的球状物体。地球可以近似地看作一个球体，但准确地说它实际上只是一个椭圆球体(参见**椭圆球体**)。

**标准纬线 (Standard parallel)**——制图学投影中用来定义y坐标原点的纬度线。

**标准 (Standards)**——在运算中，由某个机构建立的，用来确定准确度要求、数据交换格式和软硬件系统的一套准则或规范。

**结构查询语言 (Structured Query Language(SQL))**——在关系型数据库管理系统中用来定义、处理和抽取数据的一种标准语法软件。

**面 (Surface)**——常用来描述地理信息系统光栅数据或图像数据的一个术语。它可以描述连续性的、平稳变化的现象(例如高度或温度)。甚至人口密度有时也可以被描绘成一个光栅面。

**符号 (Symbols)**——制图学上用来代表地图特征的图案元素。符号类型有点、线以及某种形态的多边形。符号化系指对图解变量(如形状、尺寸、颜色、型式和结构)的选择。

**表格 (Table)**——系指数据库管理系统中一套按行(记录或实例)和列(信息域或项目)排列的数据集。列的数量通常是按照该表结构确定的固定值，而行的数量则是可变的。

**标记图像文件格式 (Tag Image File Format(TIFF))**——这是一种标准的图像或光栅文件格式，它可以存储经过压缩或未经压缩的黑白图像、灰度图像或彩色图像。可以创建图像数据的扫描仪和其他装置常用标记图像文件格式输出。在地理信息系统中，Geo-TIFF格式被作为标准的TIFF图像文件，对遥感图像、数字正射照片或光栅地理信息系统数据集进行描述。它包括一个扩展名为.tfw的相关文件，这个文件中含有图像的地理基准信息、实际单位单元格尺寸以及其他相关信息。

**模板 (Template)**——制图学上系指地图的周边元件(边界、框线、指北箭头)的标准格式，可用于一系列标准地图。在数据库管理系统中指为多种用途而建立的空白表格，此种表格中只定义了域或条目。

**专题层 (Thematic layer)**——见层。

**专题地图 (Thematic map)**——表达一个专门概念、题材或题目的地图。专题地图可以显示定量或定性信息。

**专题 (Theme)**——系指地理信息系统中通常属于同一个题材组(如道路或住宅区)的一套地理目标，并且这套目标存储在同一个地理信息系统数据库中。

**拓扑集成地理编码参照系 (Topologically Integrated Geographic Encoding and Referencing(TIGER))**——由美国人口普查局开发，用以支持人口普查项目和调查的一种数据格式。该系统的文件是以内部格式出现并含有沿道路网络线、人口普查区段和区块边界分布的一系列街道地址的地理信息系统数据集。这个系统是为国家创建一个完整的数字化人口普查地理信息系统数据库的工作中最先取得的成果之一。

**片区(Tile)**——在地理信息系统中有时用来表示存储在独立文件中但依次邻接的数字图幅的术语。片区的形状可能是规则的(如正方形或长方形)，也可能具有不规则边界(如地区或省的边界)。将所有的片区存储在同一个地理参照系中就可以将相邻的片区暂时或永久性连接起来。

**地形图 (Topographic map)**——一般由真实地理特征(如高度等值线、河流、道路、住宅区、地标等)构成的地图。由国家测绘机构按各种比例尺绘制的标准地图一般都是地形图。

**拓扑结构 (Topology)**——在地理信息系统中用来表示地理特征(如点、线、节点、多边形)之间空间关系的术语。拓扑结构的数据库不仅可以存储单个地理特征，而且可以存储某些地理特征与其类型相同或不同的另一些地理特征之间的关系。例如，除了存储代表道路网络的一组曲线外，该系统还可存储定义道路交叉处的节点，因此可用该系统来确定沿几条不同路段伸展的路线。再如，拓扑机构的地理信息系统不是将多边形作为封闭的环线存储(作为封闭的环线存储时需将相邻多边形的边界存储两次)，而是将该条曲线只存储一次，同时将多边形在该曲线左右两侧的位置信息一道存储起来。这样做可以避免重复，有利地发挥许多地理信息系统和空间分析的功能。

**转换 (Transformation)**——通过编译、循环和换算将数字空间数据从一个坐标系向另一个坐标系的转化过程。转换可用来将数字化的数字图数据从数字化工具使用的单位(如厘米或英寸)换算成与源图的地图投影和坐标系一致的真实单位(如米或英尺)。参见**地理基准**。

**网络传输控制协议 (Transmission Control Protocol(TCP))**——作为因特网建立基础的协议之一。

**不到位 (Undershoot)**——在数字化过程中，系指没有将一条线一直延长到它本应该与另一条线相交的交点而中途停止的现象。

**通用横向墨卡托投影 (Universal Transverse Mercator(UTM))**——系指一种柱面地图投影，常用于大比例尺(即局部)的地图绘制。

**用户段 (User segment)**——全球定位系统中配备有各种类型全球定位系统信号接收器的那一个区段。

**向量数据 (Vector data)**——用点、线、面来代表目标位置和形状的一种地理信息系统数据模型，这些点、线、面基本上由x, y坐标构成。

**向量积格式 (Vector product format(VPF))**——由美国国家地图与影像局(前国防部测绘局)开发的一种向量地理信息系统格式。开发这种地理信息系统格式是想使其成为普遍接受的向量数据交换格式。

**极点 (Vertex)**——定义一条线的一系列 $x, y$ 坐标之一。线的第一个和最后一个极点通常称为节点。

**纵向集成 (Vertical integration)**——见**集成**。

**广域网 (Wide Area Network(WAN))**——借助高速通信链路或卫星将远距离计算机连接起来的计算机网络。

**万维网 (World Wide Web(WWW))**——由驻瑞士的欧洲粒子物理实验室开发的, 用于分配电子文件的系统。这些电子文件由许多位于世界各地的不同格式的文件组成, 或针对这些不同格式的文件进行编制。电子文件按照标准化的超文本标记语言(html)生成, 在用户计算机上借助网络浏览器解译这种超文本标记语言。超文本文件的位置可通过链路或被所谓的“全球资源定位符”(URLs)的地址确定。万维网取得了飞速发展, 正在变成文件和数据传输的一个重要渠道。专业地理信息系统软件使一个机构能够在万维网上提供地图服务。例如, 远程用户可以利用位于发布机构网络服务器上的地理信息系统数据库设计和显示专题地图。

另外还有一些专业术语和词汇可以从Padmanabhan和其他人(1992年)、美国土木工程师学会(1994年)、McDonnel和Kemp(1995年)以及Dent(1998年)等文献中查到。网上资源包括:

Canada Centre for Remote Sensing	<a href="http://www.ccrs.nrcan.gc.ca/">http://www.ccrs.nrcan.gc.ca/</a>
Geographer's Craft Project (University of Texas)	<a href="http://www.colourado.edu/geography/gcraft/contents.html">http://www.colourado.edu/geography/gcraft/contents.html</a>
GPS World magazine	<a href="http://www.gpsworld.com/gpsworld/static/staticH-tml.jsp?id=8000&amp;searchString=glossary">http://www.gpsworld.com/gpsworld/static/staticH-tml.jsp?id=8000&amp;searchString=glossary</a>
Perry-Castañeda Library, University of Texas	<a href="http://www.lib.utexas.edu/Libs/PCL/Map_collection/glossary.html">www.lib.utexas.edu/Libs/PCL/Map_collection/glossary.html</a>
United States Census Bureau	<a href="http://www.census.gov/dmd/www/glossary.html">www.census.gov/dmd/www/glossary.html</a>
United States Geological Survey	<a href="http://interactive2.usgs.gov/learningweb/explorer/geoglossary.htm">http://interactive2.usgs.gov/learningweb/explorer/geoglossary.htm</a>



## 附件七

### 有用的网址和URLs

#### 地理信息系统程序

Autodesk Inc.	San Rafael, CA	AutoCAD	www.autodesk.com
Bentley Systems Inc.	Huntsville, AL	MicroStation	www.bentley.com
ESRI, Inc.	Redlands, CA	ArcGIS, ArcInfo, ArcView, ArcExplorer, Atlas GIS	www.esri.com
Intergraph	Huntsville, AL	GeoMedia	www.intergraph.com
MapInfo Corp.	Troy, NY	MapInfo GIS	
Microsoft Corp.	Redmond, WA	MapPoint	www.microsoft.com
Oracle Corp.	Redwood Shores, CA	Oracle Spatial	www.oracle.com
UNSD Software Project	New York, NY	PopMap	www.un.org/Depts/unsd/softproj/index.htm
Siemens	Munich, Germany	SICAD Spatial Desktop	www.siemens.com
Smallworld Systems Inc.	Englewood, CO		
PCI Geomatics Group	Richmond Hill, Ontario, Canada	SPANS and PAMAP	www.pci.on.ca
ThinkSpace Inc.	London, Ontario, Canada	MFWorks	www.thinkspace.com
Vision* Solutions	Ottawa, Ontario, Canada	Vision*	

#### 专业软件

Blue Marble Geographics	Gardiner, ME	Coordinate management and GIS development tools	www.bluemarblegeo.com
Caliper Corp.	Newton MA	Maptitude, GIS+, TransCAD	www.caliper.com
Core Software Technology	Pasadena, CA	TerraSoar (distributed geospatial databases), ImageNet (online geospatial data distribution)	www.coresw.com
Quantum GIS		Open source software	(http://qgis.org)
Thuban		Open source software	http://thuban.intevation.org
Open EV		Open source software	http://openev.sourceforge.net

#### 遥感图像处理系统

Leica GeoSystems/Erdas	Atlanta, GA	ERDAS Imagine	www.erdas.com
Earth Resource Mapping	San Diego, CA	ER Mapper	www.ermapper.com
Clark Labs	Worcester, MA	Idrisi GIS	www.clarklabs.org
MicrolImages Inc.	Lincoln, NE	TNTmips	www.microimages.com
PCI Geomatics Group	Richmond Hill, Ontario, Canada	EASI/PACE, OrthoEngine	www.pci.on.ca
Research Systems Inc	Boulder, CO	ENVI visualization software	www.rsinc.com

### 高分辨率卫星图像和正射摄影

GeoEye	Thornton, CO	Carterra and Ikonos satellites	<a href="http://www.spaceimaging.com">www.spaceimaging.com</a>
Digital Globe	Longmont, CO	QuickBird and EarlyBird satellites	<a href="http://www.digitalglobe.com">www.digitalglobe.com</a>
Orbital Imaging Corp.	Dulles, VA	Orbimage satellites	<a href="http://www.orbimage.com">www.orbimage.com</a>
EROS Data Center	Sioux Falls, SD		
Spot Image		Spot satellites	<a href="http://www.spot.com">www.spot.com</a>
Maps Geosystems	Munich, Germany	Aerial surveys (Africa, Middle East)	<a href="http://www.maps-geosystems.com">www.maps-geosystems.com</a>
EarthSat	Rockville, MD	Satellite and mapping services	<a href="http://www.earthsat.com">www.earthsat.com</a>

### 全球定位系统

Magellan Corp.	Santa Clara, CA		<a href="http://www.magellangps.com">www.magellangps.com</a>
Ashtech	Santa Clara, CA		<a href="http://www.ashtech.com">www.ashtech.com</a>
NovAtel Inc.	Calgary, Alberta, Canada		<a href="http://www.novatel.ca">www.novatel.ca</a>
Sokkia Corp.	Overland Park, KA		<a href="http://www.sokkia.com">www.sokkia.com</a>
Trimble Navigation Ltd.	Sunnyvale, CA		<a href="http://www.trimble.com">www.trimble.com</a>
Garmin			

### 期刊

GeoWorld, GeoAsia, GeoEurope, GeoInformation Africa, Mapping Awareness, Business Geographics	GeoWorld, Fort Collins, CO		<a href="http://www.geoplance.com">www.geoplance.com</a>
GPS World			<a href="http://www.gpsworld.com">www.gpsworld.com</a>
International Journal of Geographical Information Science	Taylor & Francis, London, United Kingdom		
GeoInfosystems	Advanstar Pub., Eugene, OR		
Journal of the Urban and Regional Information Systems Association	URISA, Park Ridge, IL		<a href="http://www.urisa.org/">www.urisa.org/</a>
ISPRS Journal of Photogrammetry & Remote Sensing			<a href="http://www.itc.nl/isprsjournal">www.itc.nl/isprsjournal</a>

### 其他

National Center for Geographic Information and Analysis	Santa Barbara, CA	GIS research center	<a href="http://www.ncgia.ucsb.edu">www.ncgia.ucsb.edu</a>
International Institute for Aerospace Survey and Earth Sciences (ITC)	Enschede, Netherlands	GIS Training Courses	<a href="http://www.itc.nl/">http://www.itc.nl/</a>
European Umbrella Organization for Geographic Information (EUROGI)	Netherlands		<a href="http://www.eurogi.org">www.eurogi.org</a>
U.S. Federal Geographic Data Committee	Reston, VA		<a href="http://www.fgdc.gov">www.fgdc.gov</a>
Permanent Committee on GIS Infrastructure for Asia & the Pacific			<a href="http://www.permcom.apgis.gov.au/">www.permcom.apgis.gov.au/</a>





