지역정보시스템 (GIS)

1990. 11

경제기획원 조사통계국

이 교재의 원고는 서울대학교 사회과학대학 지리학과 유 근 배 교수(지리학 박사)가 작성하여 제공한 것입니다.

목 차

1. 서 론

최근에 들어서, 지리정보 시스템(Geographic Information System, GIS)은 세계 각국의 정부기관, 개인 기업체, 정보산업계, 그리고 관련 학술분야 등에서 큰 관심을 보이는 분야가 되었다. GIS를 주제로 하는 각종 학술회의가 계속 열리고 있으며 그 반응 또한 대단하여, 매번 참석인원이 예상치를 초과하고 있다. 미국의 텍사스주 샌 안토니오에서 1988년 11월에 열린 "국제 GIS회의(GIS/LIS '88)"는 미국 원격탐사학회, 미국 지리학회, 도시 및 지역GIS학회 등 4개의 유력한 학술단체에 의해 공동으로 진행되었다. "세계에의 접근(Accessing the World)" 이라는 표어를 내세운 이 학술세미나에는 40여개국에서 3,000여명이 참석하였으며, 200여편의 논문이 발표되었다.

일본의 경우에는 1986년의 제1회 GIS학술세미나에 참석 예정인원이 100여명이었으나 세미나 당일 참석한 사람은 500여명에 이르렀기 때문에 세미나 장소를 더 넓은 곳으로 옮기지 않을 수 없었으며, 우리나라의 경우에도 금년 6월 서울대학교 국토문제연구소 주최로 열린 GIS워크샵에 참석한 인원은 예상을 훨씬 초과한 200여명에 달했다. GIS를 제작하고 있는 회사가운데 가장 유력해지고 있는 미국의 ESRI사에서는 10년전부터 자사 제품인 ARC/INFO의 USER CONFERENCE를 개최하고 있는데, 1981년 제1회때에는 16명이 참석하였으나 제10회였던 금년 5월에는 약 1,800여명이 모였다. 이제 세계 어느나라 할 것 없이 각국 정부기관, 정보산업계, 유관 학문분야등 사회 각 분야에서 GIS에 대해 관심을 가지는 사람의 수는 급격히 증가하고 있다.

그렇다면 GIS가 최근 몇 년 사이에 이와같이 각광을 받는 분야가 될 수 있었던 이유는 무엇인가? 우선 그 이유를 찾아보도록 하자.

전 세계적으로 급속한 산업화와 도시화를 겪으면서 지역간의 불균형, 도시인구의 폭증, 환경오염, 도시교통의 혼잡, 토지이용의 부적절 등 많은 문제를 안게 되었다. 또한, 정보사회로 진입하면서 정부기관을 비롯한 각종 기관은 정보공개에 대한 요구에 직면하게 되었다. 이러한 문제를 해결하기 위해서는 정확하고도 시기적절한 정보가 뒷받침되어야 하지만, 자료의 양이 방대해지고 종류가 다양해짐에 따라 전통적인 자료관리방법이나 기술로는 감당할 수 없게 되었다.

대부분의 국토정보는 현상의 장소·설명(속성)·시간으로 정의되는 지리정보(Geographic Information)로 구성되어 있다. 연구에 따르면, 미국 정부 각 부문의 정책결정과정에서 80%이상이 지리자료를 사용한다고 한다. 정부와 기업계의 입장에서 볼때, GIS를 통해 정부나 기업의 기구는 각각 담당하는 지역의 각종 현황을 정확히 파악함으로써 당면한 제반 문제들을 해결할 수 있고, 업무수행의 효율성을 제고시킬 수 있다.

또한, GIS는 급속히 발전하는 고도의 과학기술분야로서 유용성이 매우 높기 때문에 학술적으로도 급속한 진보를 보이고 있다. GIS는 지리학, 지질학, 자원공학, 임학, 해양학, 도시계획학, 지역계획학, 조경학, 토목공학, 경영학, 부동산 등과 같은 공간정보와 관련된 여러 학문분야에서 매우 효율적인 연구도구(Research tool)로 활용되고 있다.

한편, 정보산업체에서 볼 때, 그 시장규모가 연간 수억 달러에 달하며 매년 40%내외의 성장을 기록하는 GIS야말로 매력적인 분야가 아닐 수 없다. 따라서, 세계적인 컴퓨터회사들(IBM, DEC, PRIME, HP, SUN, APOLO, APPLE등)은 연구개발비용을 엄청나게 투입하면서까지 GIS시장에서 자사의 시장점유율을 확보하려고 노력하고 있다. GIS소프트웨어시장의 경우도 하드웨어와 마찬가지로 소프트웨어의 성능을 향상시켜 기술적인 측면

에서 우위를 점하려고 전력을 다하고 있다.

이와같은 정부나 기업체, 학계 그리고 정보산업체 전문인력의 움직임 속에서, GIS는 이 시대의 과학기술혁명에 있어서 심원한 획을 그을 수 있는 영향력있는 한 분야가 되었다.

2. GIS의 정의 및 자료

2.1 지리학과 정보시스템

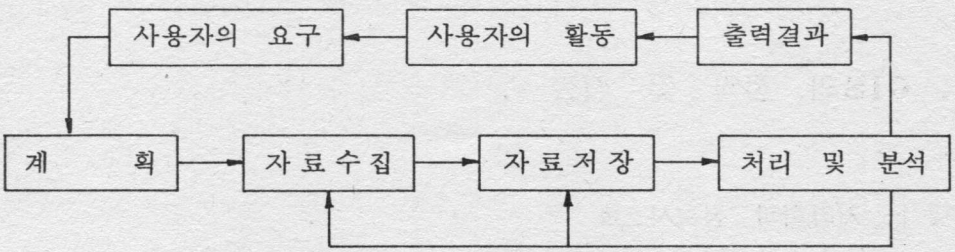
독일의 철학자이며 자연지리학자인 칸트는 학문을 크게세가지로 나누었다. 첫째, 특정한 대상, 군락, 또는 현상을 다루는 학문으로 생물학, 임학 등을 들었다. 두번째로, 시간의 경과에 따른 현상의 변화를 다루는 학문으로 역사학을 들었으며, 세번째로는 공간적 맥락(Spatial context)에서 자연현상 및 인문현상을 다루는 지리학을 거론했다. 전통적으로 지리학은 사람과 땅과의 관계에 중점을 두어왔고, 이 관계를 효율적으로 표현하는 지도가 key tool로 사용되어 왔다.

한편, 정보 시스템(Information System)이란 현상을 관측하고, 자료를 수집·저장·분석하여 의사결정에 유용한 정보로 압축하는 일련의 처리과정이라고 정의할 수 있으며, 자료원에서 정보의 사용까지를 다루는 end-to-end system이다〔그림 1 〕.

한편, 지도는 공간자료를 수집·분석하고 유용한 정보를 선별하여 지리좌표의 틀속에 정리하고 있으며, 정보를 명석 판명하게 전달하는 수단이기 때문에 정보시스템의 좋은 예라고 할 수 있다.

GIS는 공간좌표 또는 지리좌표(Spatial or Geographical Coordinate)에 참조된 자료와 이와 연관된 비공간자료를 다루는 것으로,"모든 형

태의 지리정보를 효율적으로 수집·저장·갱신·처리·분석 표시하기 위해 구축한 하드웨어 소프트웨어 지리자료 인적자원의 조직체"라고 정의할 수 있다. GIS는 서로 다른 축적(Scale)·시간(Time)·형태(Format)의공간자료를 연관시키고 통합시키는 수단이다.



〔그림1〕 정보 시스템의 개요

인류가 가지고 있는 지구에 대한 정보는 극히 한정되어 있고, 제한된 정보조차도 서로 긴밀하게 연관되어 있다고 볼 수 없다. 사람들은 현미경을 사용하여 분자수준의 현상을 관찰하기도 하고, 기상현상은 현상자체가 너무 크기 때문에 정지기상위성을 통하여 관측한다. 이 두가지 극단적인 축척(Scale)의 지구정보는 인간의 인식구조에서 얼마나 연관되어 있는가? 정보와 정보가 무한하게 연관되어 새로운 정보가 무한하게 창출된다면 지표공간은 어떻게 변모할 것인가? 공간정보를 처음으로 보편화시켰던 국세조사가 가져왔던 혁신을 생각한다면 그 내용과 범위를 상상할 수 있을 것이다.

GIS는 최근에 발달된 과학 기술 부문이기 때문에 아직 명확하게 정리되지 않은 부분이 많고, 처리·분석 및 출력 기능이 다양하고 적용되는 분야에 따라 강조되는 기능이 다르기 때문에 컴퓨터 지도학, 컴퓨터 응용 디자인(CAD), 원격 탐사 등의 시스템과 혼동하고 있는 사람들이 많다.

GIS와 컴퓨터 지도학은 많은 공통점이 있으나, 몇가지 점에서 큰 차이

를 보인다. 우선 공간자료의 변환에서 GIS에서는 위상관계(Topology), 공간적 위치, 비공간적 속성, 위치와 속성의 결합 등 폭넓은 기능을 가지고 있다. 또한 컴퓨터 지도학에서 비공간적 자료란 색, 선의 형태, 부호등을 가리키는데 반해, GIS에서는 토지이용의 종류, 소유권, 도로의 폭·차선·교통량, 식생의 종류 등을 말한다.

CAD에서는 수학적으로 정의되는 직선·곡선·도형으로 사상(Entity)을 표현하고, 사상의 Parent/Child의 관계를 이용하여 Object-oriented data base를 구축한다. 이에 반해 지리정보의 도형은 수학적으로 간명하게 정의될 수 없고, 자료의 양이 많고 복잡하여 Parent/Child의 관계로 조직될 수 없다.

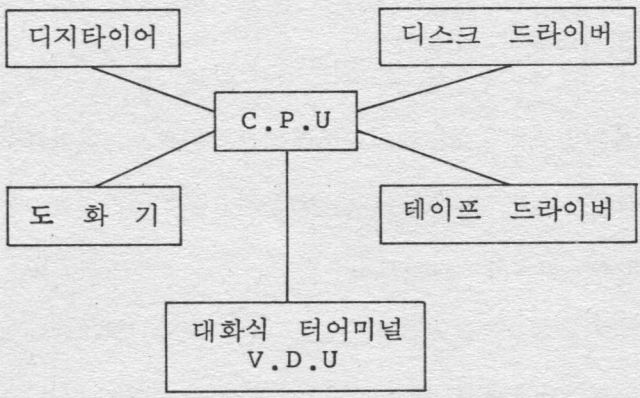
2.2 지리정보시스템의 구성요소

지리정보시스템은 하드웨어, 소프트웨어 모듈, 그리고 적절한 시스템 디자인의 세가지 요소로 이루어진다. GIS가 좋은 기능을 발휘하기 위해서는 세가지 요소가 균형을 이루어야 한다.

기능적 측면에서 GIS는 자료 획득(Data Acquisition), 전처리(Preprocessing), 자료관리(Data Management), 처리 및 분석(Manipulation & Analysis), 출력(Product Generation)의 다섯가지 기본기능으로 구성된다. 적용분야에 따라 강조되고 확장되는 요소가 다르지만, 이들 요소는 서로 연관된 것으로 연속적인 과정으로 파악해야 한다. 기능은 다음 장에서 좀 더 살펴보기로 하자.

지리정보 시스템의 일반적인 하드웨어 구성은 〔그림 2〕와 같다. 컴퓨터 혹은 CPU(Central Processing Unit)가 자료와 프로그램이 저장된 디스크 드라이브(Disk Drive)에 연결되어 있다. 디지타이저(Digitizer)

나 다른 주변 기기는 지도와 문서에서 추출된 자료를 숫자의 형태로 변환하여 컴퓨터에 보낸다. 도화기 (Plotter) 나 다른 출력기기는 자료처리 결과를 출력하는데 이용되며, 테이프 드라이브 (Tape Drive) 는 자료나 프로그램을 테이프에 저장하거나 다른 시스템과의 통신에 이용된다. 컴퓨터간의 통신은 특별한 자료선 (Data Line) 혹은 '모뎀 (Modem : Modulation —demodulation) '이라는 기기와 전화선으로 구성된 네트워어크 시스템 (Network system) 으로 이루어진다. 사용자는 컴퓨터와 주변기기들을 터미널의 화면 (Visual Display Unit : VDU) 을 통해 조작한다. 터미널에는 마이크로 컴퓨터 또는 지도를 신속하게 보여주는 특수한 하드웨어가 이용된다.



〔그림 2〕 지리정보 시스템의 주요 하드웨어 구성

2.3 GIS의 자료

GIS에서 처리하는 대상은 지리자료 (공간자료) 이다. 공간적인 분포의 특성을 가지는 것이라면 모두 지리자료라고 할 수 있으며, 지도상에 이러한 각종 지리자료를 기입하여 왔다. 흔히 사용되는 주제도로는 지형도, 지질도, 토양도, 지적도, 그리고 지하시설물 (상하수도, 지하 전화배선, 지하 가

스관 등)을 표시한 지도 등이 있다. 지리자료는 부호로 표시되는데, 그밖에 주소나 발생건수 등은 글과 숫자로 기재되기도 한다. 완벽한 지리자료가 되기 위해서는 공간자료(Spatial Data)와 속성자료(Descriptive or Attribute Data)가 함께 정리되어야 한다. 공간자료는 각종 지리현상의 위치와 형상 및 사상간의 공간상 상대적 위치관계를 말하는데, 지도상에서는 점, 선, 면을 사용하여 표시된다. 속성자료는 점, 선, 면으로 표시된 각종 지표공간의 특성을 말한다. 예를들어, 하나의 점으로 표시된 어떤 지표공간을 수질관측소라고 하면, 한쌍의(x , y)좌표로 그 위치가 공간자료로 변환한다. 속성자료에는 관측소의 명칭, 기호, 특정한 시각에 측정한 생화학적 산소 요구량(BOD)이나 용존 산소량(DO)등이 포함된다. 도로와 같은 선자료(線資料)는 일련의 점을 상호연관시켜 나타낸다. 속성자료인 도로의 명칭, 노폭, 노면재료, 교통량 등은 문자와 숫자로 데이타 베이스에 기재된다. 다각형 또는 면자료(面資料)인 행정구역도와 토양도 등은 하나 또는 여러개의 선분으로 구성되는 폐곡선으로 공간자료가 만들어지고, 속성자료로는 폐곡선으로 구분되어진 지역의 인구특성이나 토양의 성질 등이 기록된다.

지리자료 가운데 일부는 지형·토양·지질 등과 같이 그 속성이 쉽게 변하지 않는다. 그러나, 지적(地籍), 토지이용, 인구자료, 수질상황 등과 같은 자료는 수시로 변할 수 있는 것들이다. 따라서, 데이타 베이스의 내용을 어떻게 유지하고 갱신하는가 하는 것은 GIS의 구축과 발전에 있어서 반드시 고려해야 할 중요한 문제이다.

지리자료의 대부분은 정부기관이 수집·관리하는데 종류에 따라 수집과 관리하는 기관이 서로 다른 것이 일반적이다. 이와같이 행정 혹은 관료조직에 의존하는 지리자료의 수집 및 관리 방식은 관리의 측면에서는 매우 효율적이긴 하지만 타 기관에서 관리하는 지리자료의 활용이나 여러 기관에

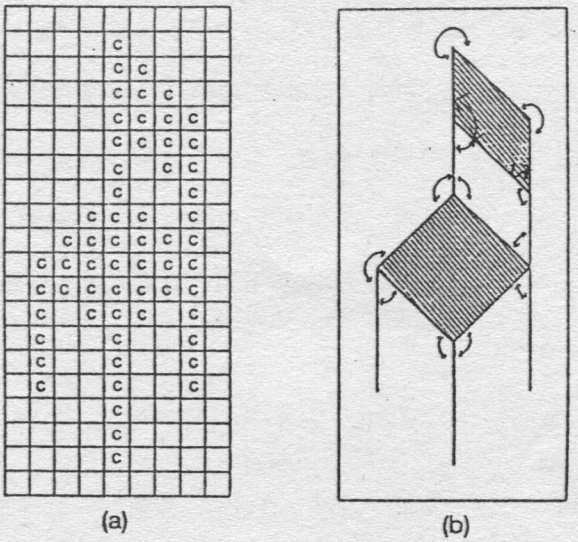
서 제각기 관리하는 여러 분야의 지리자료를 정리하고 통일시키는 데에는 어려움이 많다. 그 가운데 가장 심각한 것은 자료의 수집·분류·정리 등에 있어서의 표준이다. 이 문제는 동남아 각국이 모두 당면하고 있으며, 구미의 몇몇 선진국에서도 최근에서야 표준을 완성했다. 한편, GIS의 기술적인 발전은 이러한 문제의 해결을 용이하게 할 것이다.

2.4 GIS의 자료구조

사람의 눈은 형태를 인식하는데 매우 효과적이다. 그러나, 컴퓨터는 공간 패턴을 다루고 표시하는 방법을 정확히 지시받아야 한다. 공간정보를 표현하는 데에는 대조적이면서도 상호보완적인 두가지 방법, 즉 명시적(Explicit)방법과 내재적(Implicit)방법이 있다. 〔그림3〕은 컴퓨터에서 의자를 두가지 표현방법으로 나타낸 것이다. 명시적 표현이란 의자의 형태를 그리드(Grid)나 래스터(Raster)구조에서 점의 집합으로 나타낸 것을 뜻한다. 점의 집합으로써 책상이 아닌 의자를 표현하였으며 각 셀(Cell)은 'C'라는 코드값을 똑같이 가지고 있다. 실제로 'C'라는 것은 그 자체로 나타나지는 않고 수치나 색깔, 명암도 등으로 표현될 것이다. 의자에 관한 간단한 자료구조는 다음과 같이 볼 수 있다.

의자 ------ 기호화/채색화 ------ 셀X

내재적 표현은 시작점과 끝점, 그리고 연결의 형태로 표시된 선의 집합을 이용한다. 선의 시작점과 끝점은 의자의 형태를 표현하는 벡터(Vector)를 정의한다. 선과 선 사이의 포인터는 의자의 형태를 만들기 위해 선이 어떻게 연결되어 있는가를 컴퓨터에 알려준다. 자료구조는 다음과 같다.



〔그림 3〕 래스터(a)와 벡터구조(b)

〔그림3〕에 표현된 두가지 표현방법은 여러가지 점에서 차이를 보이고 있다. 첫째로 내재적 표현은 자료의 양이 적다. 즉 정보의 저장에 쓰이는 공간이 매우 적다. 벡터 표현은 11개의 좌표쌍과 14개의 연결포인터를 사용하는데 비해 래스터표현은 60개의 셀을 사용한다. 둘째로 벡터 표현은 외관상으로 래스터 표현보다 훨씬 산뜻하다. 〔그림 3〕의 경우 래스터 영상(a)으로 (b)와 같은 해상력을 가지려면 0.5㎜ 크기의 격자를 만들어야 하므로 이는 470개의 XY좌표가 필요하다. 세번째로 연결성의 정보(Connectivity Information)는 직접 공간검색에 이용될 수 있다. 그러나, 의자의 형태나 크기가 변하면 래스터식 표현은 벡터식보다 훨씬 빠르고 쉽게 고쳐질 수 있다. 래스터식 표현에서 정보의 수정은 단지 어떤 값을 지우고 새로운 값을 쓰기만 하면 된다. 벡터식 표현에서는 좌표값이

수정되어야 할 뿐 아니라 연결성도 다시 조직되어야 한다.

위상적 정보를 나타내는 가장 근본적인 이 두가지 방법은 다음과 같이 요약될 수 있다.

o 래스터 구조 : 좌표에 의해 정의된 셀 (Cell)의 집합. 각 셀은 속성값에 따라 독립적으로 기록된다.

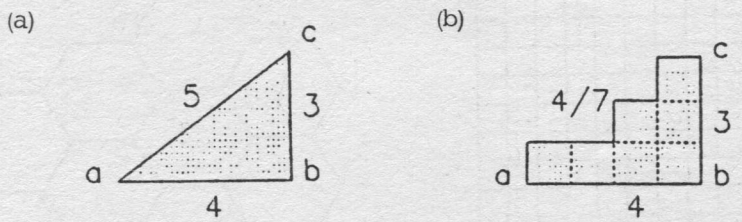
o 벡터 구조 : 점, 선, 면의 세가지가 있다. 점은 셀과 비슷하나 면적은 없다. 선과 면은 서로 연결되는 XY 좌표의 집합이며, 좌표는 주어진 속성과 관련시킬 수 있다.

2.4.1 래스터 자료구조

래스터 자료구조의 가장 간단한 형태는 그리드 (Grid), 셀 (Cell), 또는 픽셀 (Pixel)로 구성된 배열 (Array, Raster, Matrix or Lattice)이다. 각 셀은 행과 열의 값으로 참조되며, 지도화되는 속성의 값이나 유형을 나타내는 수치를 가지고 있다. 래스터 구조에서 점은 하나의 셀로 표현된다. 또한, 선은 어떤 방향으로 배열되어 인접하고 있는 셀들에 의해 구성 (표현)되며, 면은 사방으로 인접하고 있는 셀의 집합으로 표현된다. 이 자료구조는 지리정보를 표현한 2차원 자료공간을 연속적인 것이 아니고 분할된 (Quantized)된 것으로 가정한다. 따라서, 표현되는 사상에 비하여 셀의 크기가 클 때 길이와 면적의 계산에 큰 영향을 준다. 예를 들면 그림 4(a)에서 a와 c사이의 거리는 5이다. 반면 그림 4(b)에서 a와 c의 거리는 이것이 통과하는 셀의 변을 세는가 또는 셀을 세는가에 따라 7이나 4가 될 수 있다. 그림 4(a)의 면적은 6이며 그림 4(b)의 면적은 7이다. 셀의 크기에 따라서 정확도가 달라지기 때문에 수치화상처리 (Digital Image Processing)와 같은 분야에서

는 분할화된 표면을 연속적인 것으로 간주하여 미분 가능 함수를 적용하고 있다.

래스터 표현은 지리공간을 평평한 데카르트면으로 가정하며, 각 셀은 묵시적으로 정사각형의 지면을 나타낸다고 가정한다. 해상력 (Resolution)과 축적은 데이타베이스내에서 차지하는 셀의 크기와 셀이 실제 지표상에서 차지하는 면적의 관계에 따라 정해진다.



〔 그림 4 〕 래스터화가 면적과 거리에 미치는 영향

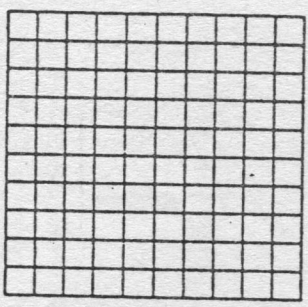
2.4.2 공간 분할 기법

평면을 같은 크기의 셀로 나누는 방법 (Tessellation) 에는 여러가지가 있다〔그림 5 〕. 가장 많이 사용되는 것으로는 정사각형이 있는데, 도트 매트릭스 프린터가 일반화되기 이전에는 라인프린터의 활자 형태에 따라 직사각형도 많이 사용되었다.

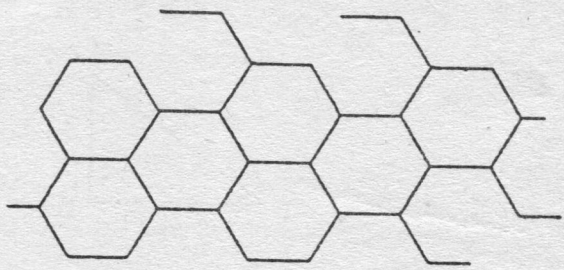
사각형 격자는 행렬로 간명하게 정리될 수 있다는 큰 장점이 있는 반면에, 인접격자에 대한 정의에 있어서 문제가 있다. 즉, 격자의 변에 인접한 4개의 셀을 인접격자로 보느냐 대각선 방향의 4격자도 포함시켜 8개로 보느냐 하는 것이다. 격자사이의 거리도 1과 1.41이 되기 때문에 또 다른 문제가 된다.

이에 대한 대안으로 정육각형법과 정삼각형법이 있다. 정육각형법은 인접격자의 정의가 명백하며 그 거리도 모두 같다는 장점을 갖고 있어 많은

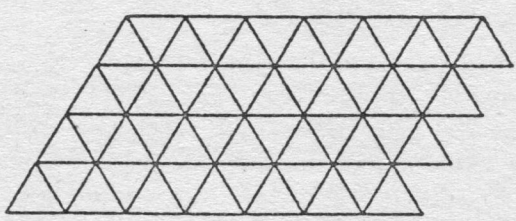
관심을 모으고 있다. 그러나 정사각형과는 달리 육각형은 같은 형태로 분할될 수 없기 때문에 해상력의 조정이 어려우며, 좌표화 또는 수치화하는 데에 부대공간 (Overhead)이 필요하다.



Rectangular



Hexagonal



Triangular

〔 그림 5 〕 공간분할기법 (**Tessellation**)

2.4.3 압축된 래스터 구조

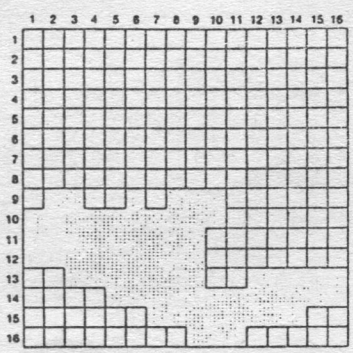
래스터 자료를 압축하는 주요방법으로는 체인 코드(chain code), 런 렝쓰 코드(run-length code), 블록 코드(block code), 그리고 사지수형(quadtree)의 네가지가 있다.

〔1〕 체인 코드

그림 6에서 경계선은 그 시작점과 동·서·남·북 방향(도엽의 위쪽을 북으로 했을 때)으로 순차 진행하는 단위 벡터에 의해 정의된다. 동은 0, 북은 1, 서는 2, 남은 3 등 숫자로 방향을 정의한다. 행 10, 열 1의 cell에서 시작한다면, 지역의 경계선은 시계방향으로 다음과 같이 입력된다.

$0, 1, 0^2, 3, 0^2, 1, 0, 3, 0, 1, 0^3, 3^2, 2, 3, 0^2, 1, 0^5, 3^2, 2^2, 3, 2^3, 3, 2^3, 1, 2^2, 1, 2^2, 1, 2^2, 1, 2^2, 1^3$

여기서 지수는 각 방향으로 중복되는 단계(cell)의 수를 표현하고 있다. 체인 코드는 압축시켜 저장하는 데에 매우 효과적인 방법일 뿐만 아니라, 면적과 둘레의 계산, 첨점(sharp turn), 오목성(concavity)등의 연산을 쉽게 할 수 있다. 반면에 합집합(union)이나 교집합(intersection)과 같은 연산은 어려우며, 또한 지역의 모든 경계선이 저장되어야 하기 때문에 자료의 중복이 불가피하다는 단점을 가진다.



〔그림 6〕 래스터 지도와 지역(region)

〔2〕 런 렝쓰 코드

각 지도단위에 속해 있는 점들을 각 열마다 왼쪽에서 오른쪽 방향으로 시작 셀과 끝 셀을 저장하는 방법이다. 〔그림 6〕은 다음과 같이 변환된다.

Row 9 2,3 6,6 8,10

Row 10 1,10

Row 11 1,9

Row 12 1,9

Row 13 3,9 12,6

Row 14 5,16

Row 15 7,14

Row 16 9,11

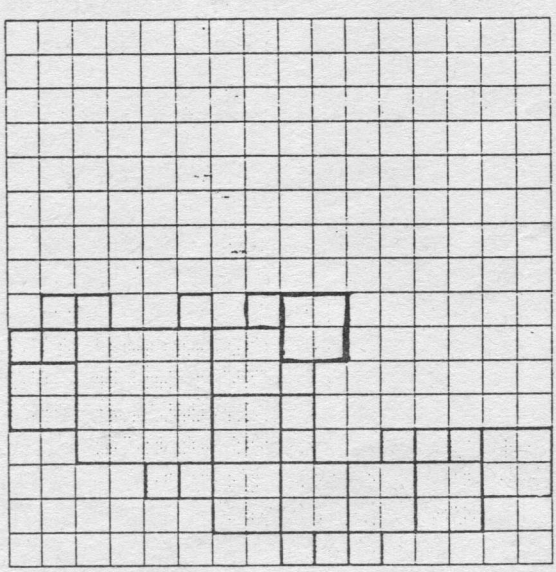
이 예에서 〔그림 6〕의 69(16×16)개의 셀이 정보의 손실없이 22개로 압축되어 입력된다. 따라서 자료를 저장하는 공간이 절약된다. 확실히 런 렝쓰 코드는 일대다수의 관계(한 지도단위가 다수의 셀로 구성된)가 있을 때에 자료저장이 용이하다. 특히 이것은 자료의 저장에 제한을 받는 소형컴퓨터에 적당하다. 그러나 너무 지나친 자료의 압축은 지도제작이나 자료조작에서 처리시간을 길게 만들 수도 있다.

〔3〕 블록 코드

런 렝쓰 코드의 착상은 지도화 되는 영역을 타일(tile)형태의 정사각형 블록을 사용함으로써 2차원으로 확대될 수 있다. 〔그림 7〕은 〔그림 6〕의 래스터 지도를 어떻게 정방형으로 표현하는가를 보여 준다. 자료구조는 세가지 숫자 즉, 원점(중간점이나 최하점)의 XY좌표와 사각형의 지름으로 되어 있다. 이것을 중간 축 변환(MAT:medial axis transform-

ation)이라고 한다. 〔그림 6 〕의 지역은 1개의 16단위 셀로 이루어진 정사각형과 9개의 4단위 정사각형, 그리고 17개의 1단위 정사각형으로 저장된다. 각 정사각형에 2개의 좌표가 필요하므로 이 지역은 57개의 숫자로 저장될 수 있다. 즉, 54개는 좌표값, 3개는 셀의 크기 또는 지름을 표시한다. 주어진 지역에서 맞춰지는 정사각형이 클수록 경계선은 더욱 단순해지기 때문에 이 방법은 효율적이다. 런 렝쓰 코드와 블록 코드 모두 크고 단순한 모양에는 효과적이나, 기본적 셀보다 불과 몇배의 크기를 갖는 지도단위로 복잡하게 구성된 경우에는 비효율적이다. 중간축 변환(MAT)은 합집합이나 교집합 연산을 하거나 지도 단위의 길이와 같은 특성을 포착하는데 장점이 있다.

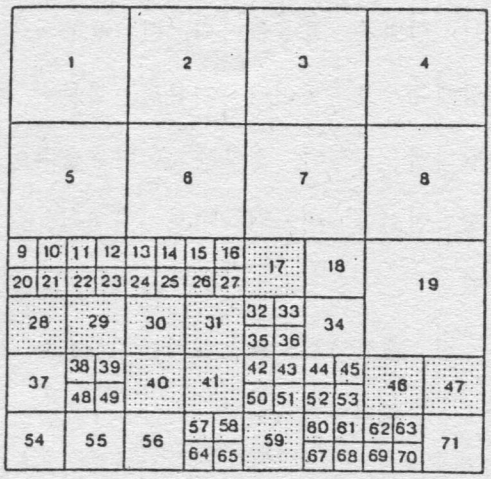
한편, 블록 코드나 런 렝쓰 코드로 저장된 자료는 몇몇 연산을 위해서 단순한 래스터 형태로 전환되어야 한다.



〔그림 7 〕 중간축 변환과 블록

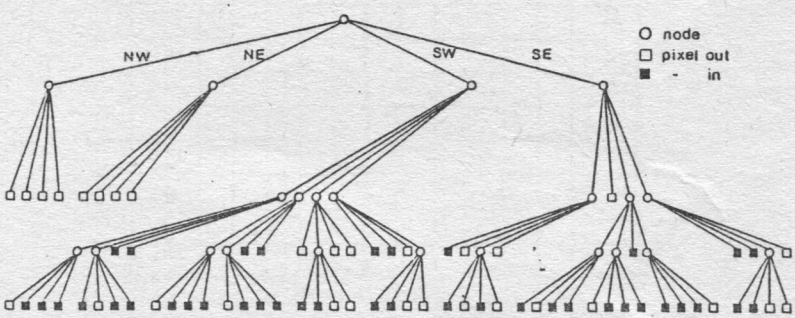
〔4〕 사지 수형 (quadtree)

자료를 축약하는 4번째 방법은 위계적으로 배열을 4분원 (quadrant)으로 나누는 것이다. 배열(지역)을 한 단계씩 밟아 사분원으로 분할시키면서 어떤 단계의 분원이 그 지역(mapping unit)에 완전히 포함되는가를 찾는다. 분할에서 가장 적은 단위는 하나의 화소이다.



〔그림 8〕 4지수형과 지역(region)

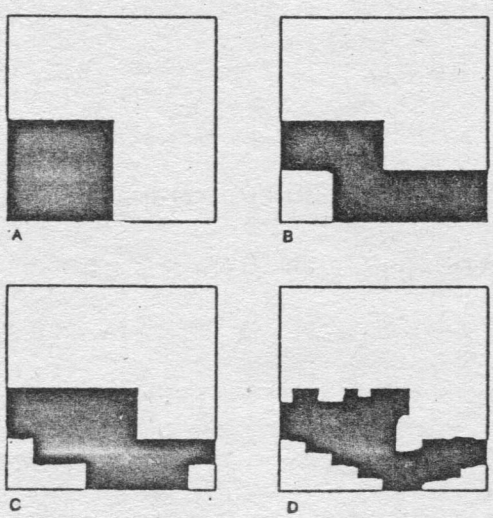
〔그림8〕에서는 〔그림6〕의 단위지역을 점진적으로 나누어 사분원 블록으로 채우고 있다. 이 블록 구조는 사지수형(quadtree)이라는 4개의 가지(degree)를 가지는 나무의 형태로 나타낼 수 있다.



〔그림 9〕 4지수형도

〔그림6〕의 4지수형은 〔그림9〕와 같다. $2^n \times 2^n$점의 배열은 4지수의 종결절(root node)이며, n의 높이(level)를 갖는다. 각 결절은 4개의 가지인 북서, 북동, 남서, 남동의 네개 분원을 갖는다. 엽결절(leaf node)은 더이상 분할되지 않은 4분원을 가리킨다.

4지수형은 다른 방법에 비해 많은 장점을 가지고 있다. 지역의 표준성(standard region property)이 효율적으로 계산된다. 4지수형은 가변적인 해상력을 가지는 배열(variable resolution array)이기 때문에 정보가 비어있는 부분을 단순하게 처리하면서도 정보가 있는 부분(면적)은 자세하게 다룰 수 있어 용량을 절약할 수 있고 검사시간을 단축시킬 수 있다〔그림 10 (a-d)〕.



〔그림 10〕 4지수형구조가 각 수준(n)에서 나타내는 지역형태

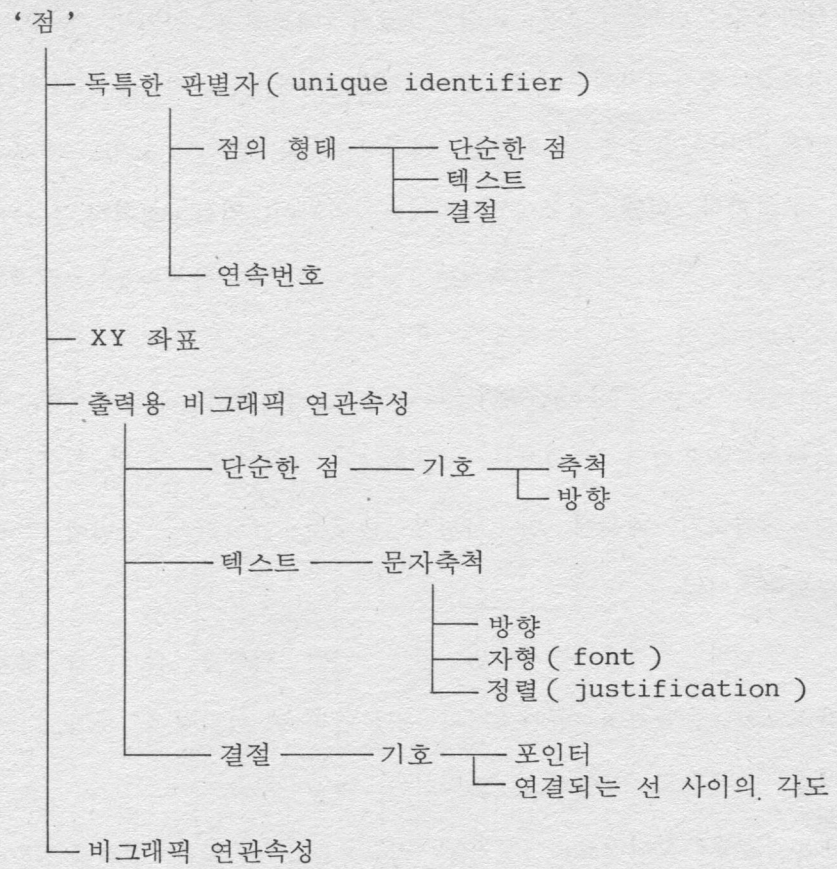
4지수형의 가장 큰 문제점은 4지수 표현이 적용마다 다른 결과가 나타날 수 있다는 것이다.(translation-variant). 즉, 같은 형태와 크기를 가진 두개의 지역이 다른 4지수형으로 될 수가 있다. 따라서 이 방

법은 형태의 분석이나 패턴의 인식에는 적합하지 않다. 그러나, 4지수형은 이질적인 지역을 쉽게 분리해 내고 도형내의 구멍(hole)을 표현할 수 있다. 이 방법은 적용범위나 개발여지가 많은 기술이며, 그 이용에 대한 관심이 점차 고조되고 있다.

2.4.4 벡터(vector)자료구조

벡터식 표현의 목적은 사상을 가능한 한 정확하게 표현하는 데에 있다. 좌표공간을, 래스터 공간과 같이 분할된 것이 아니라 위치, 길이, 차원을 정확하게 표현할 수 있는 연속적인 것으로 가정한다. 그러나 컴퓨터 용량에 제한이 있고, 또한 벡터 표현기기가 래스터 기기보다 해상력이 훨씬 좋지만 기본적 스텝(basic step size)을 내재하고 있기 때문에 엄밀한 의미에서는 연속적이 아니다. 벡터 시스템에서 지리사상은 점•선•면의 조합으로 표현된다.

점사상은 하나의 XY좌표로 위치가 정의되는 지리적 사상이라고 할 수 있다. XY좌표 이외에도 이 '점'이 무엇이며 어떤 속성을 가지고 있는가를 나타내는 정보가 추가되어야 한다. 예를 들어 '점'은 다른 정보와 관련없이 기호화 될 수 있다. 자료의 레코드는 기호와 그 크기, 기호의 방향에 관한 정보를 포함해야 한다. '점'이 문자일 때에는 자료 레코드는 문자의 철자, 활자의 형태(font), 정렬(좌측정렬, 우측정렬, 중앙정렬), 축척, 방향 등에 관한 정보뿐 아니라 다른 비그래픽 속성과의 관련 정보도 포함해야 한다. 〔그림11〕은 '점'사상 자료구조의 예를 보이고 있다.



〔그림 11〕 점 사상의 벡터자료구조

선 사상은 2개 이상의 좌표로 구성되어 있는 선형 (linear feature) 이다. 가장 단순한 선은 시작점, 끝점 (2개의 XY좌표), 그리고 기호를 지시하는 레코드로 구성되어 있다. 기호를 지시하는 파라미터 (display symbol parameter)는 그래픽 출력장치 (disply device)에 선분이나 점선을 불러낸다. '아크 (arc)', '체인 (chain)', '스트링 (string)' 은 연속적인 복잡한 선을 묘사하는 다수 (n개)의 XY좌표 집합이다. 두 개의 XY좌표로 정의되는 선분이 짧을 수록, 또는 XY좌표의 수가 많을 수록 체인은 복잡한 곡선에 가까워진다. 한편, 처리시간이 더 소요되지만

특정한 한개의 값(number)을 저장함으로써 자료저장 공간을 절약할 수 있다. 선형자료를 출력기기에 보낼 때 이값을 통해서 기기 운용 루틴(display drive routine)은 수학적인 내삽함수(B spline)가 좌표계에 적용되도록 지정한다. 이때 좌표계는 이미 기억되어 있는 것이다. '점' 또는 단순한 선과 같이 체인도 출력에 관한 정보를 지시하는 레코드와 함께 저장될 수 있다.

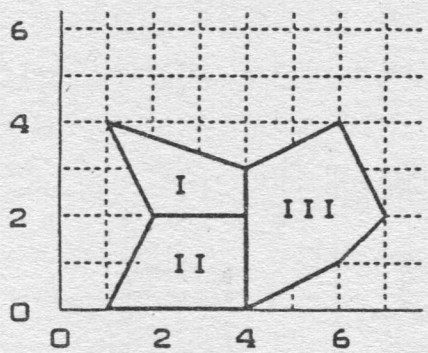
다각형(polygon 또는 region)은 벡터 데이타베이스에서 여러가지 방법으로 표현될 수 있다. 지리정보 시스템에서 대부분의 주제도는 다각형을 사용하여 제작하기 때문에 면 사상을 다루고 표현하는 방법에 많은 관심이 집중되어 있다.

다각형 자료구조의 목적은 면사상의 형태, 인접 다각형, 위계, 즉 위상적 속성을 효율적으로 운용하여 공간조직의 기본단위인 다각형이 가지고 있는 특정한 속성을 주제도로 표현하고 조작하는 데 있다.

다각형 네트워크를 지리자료화 시키는 데는 다음과 같은 전제조건이 있다. 첫째, 지도에서 단위 다각형(region)은 각각 독특한 모양, 둘레, 그리고 면적을 가진다. 여기에는 래스터 방식에서와 같이 기본적인 표준 단위(화소의 크기)가 없다. 규칙적으로 배열된 미국의 시가계획도에서 시가구역의 전부 혹은 일부가 정확히 같은 모양과 크기를 가졌다고 볼 수 없다. 토양도나 지질도에서도 지도단위(map unit)의 형태나 크기가 같은 경우는 거의 없다. 둘째, 지리적 분석을 위해서 자료구조가 하계망(stream network)에서와 같이 각각의 다각형에 인접한 다각형을 기록할 수 있어야 한다. 세째, 주제도에서 다각형은 모두 같은 수준(level)이 아니다. 호수에 있는 섬을 예로 들면 호수라는 큰 다각형안에 섬이라는 작은 다각형이 있는 경우에 두 다각형의 수준은 다른 것으로 처리된다.

〔다각형 구조〕

다각형 구조에서 각 데이타 레이어는 〔그림 12 〕와 같이 다수의 다각형으로 구성되어 있다. 다각형은 다각형을 둘러싼 선분의 좌표로 부호화 되어있고, 다각형은 각각 저장되어 있다. 이 구조에서는 참조 시스템(referencing system)이나 위상관계(topological organization)가 명시적으로 정의되어 있지 않다. 토지이용 소유권과 같이 속성자료는 좌표자료와 함께 수록될 수 있다. 이 구조에서는 두개의 인접한 다각형이 공유하고 있는 변이 두번 정의되는 등 저장공간의 이용이 비효율적이다.



Polygon I:	Polygon II:	Polygon III:
1,4	2,2	6,4
4,3	4,2	7,2
4,2	4,0	6,1
2,2	1,0	4,0
		4,2
		4,3

〔그림 12 〕 다각형 구조

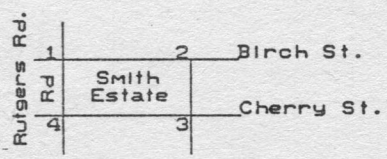
〔DIME 구조〕

명시적 위상 관계를 지리 자료구조에 처음 도입시킨 것은 미국 통계국의 DIME(Dual Independent Map Encoding) 시스템이다. DIME 데이타 화일의 기본 요소는 두개의 끝점으로 정의된 선분(line segment)이다. 복잡한 곡선은 이 선분을 연결시켜 표현한다. 각각의 선분은 양쪽 끝부분에 각

각 한개의 포인터를 가지고 결절과 연결되어 있으며, 또한 다각형을 조립하기 위한 부호(code)를 가지고 있다. 〔그림 13 〕은 간단한 DIME 구조를 나타내고 있다.

각 선분에는 선분 이름, 선분이 시작되고 끝나는 점(from node/to node), 그리고 선분의 좌우에 위치한 다각형이 저장된다. 헤더를 이용하면 속성정보가 쉽게 첨가할 수 있다. 그림에서는 헤더, 우편번호(zip code), 선분이름(가로명)을 이용하여 지번(address)을 확인할 수 있다.

DIME에는 결절에서 만나는 선분을, 혹은 하나의 선분에서 인접하는 선분을 역으로 검색(point-back)할 수 없기 때문에 복잡한 굴곡을 갖는 선분을 처리하거나 다각형의 경계선을 맞추는 것은 어렵다. 같은 가로명을 갖는 선분이 횡단선으로 단절될 때 많은 컴퓨터 시간이 소요된다. 미 통계국에서는 DIME 화일의 오차를 찾기 위해서 ARITHMICON이란 편집 프로그램을 사용하고 있다. 또한 이 자료구조에서는 복잡한 곡선을 다수의 직선으로 구성하고 있기 때문에 데이타의 양이 크다.



Header Items: Groups of Segments

Header Number	Zip Code	Area Code	Tract
1000	93106	805	14
1001	93117	805	14

Segment Codes: Each Line Segment

Segment Name	Nodes From	Nodes To	Polygon Left	Polygon Right	Addresses Left Low	Addresses Left High	Addresses Right Low	Addresses Right High	Header
Birch St	1	2	-	Smith Est	101	175	102	178	1000
Cherry St	3	4	-	Smith Est	103	177	104	180	1000
Rutgers Rd	4	1	-	Smith Est	8602	8686	8603	8685	1000

Node Locations:

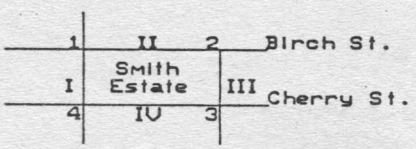
Node	Easting	Northing
1	127,251	1,340,600
2	127,352	1,040,601
3	127,350	1,040,584
4	127,256	1,040,502

〔그림 13〕 DIME 화일구조

〔아크-노드(arc-node) 구조〕

아크-노드 구조에서는 자료를 점-선-다각형의 위계조직으로 구축한다. 〔그림 14〕에서는 지도내의 노드를 x, y 좌표로 정의하고, 정의된 노드를 사용하여 아크를 정의한다. 다각형은 다시 그를 둘러싸고 있는 아크로 정의된다. 이 자료구조에서는 점·선분·다각형이 한번씩 저장되기 때문에 매우 비효율적이다. 다양한 속성정보를 쉽게 저장할 수 있고,속성정보들이 도형과 명시적으로 연결될 수 있다는 장점을 갖고 있다. 즉, 교통통제 장치에 관한 속성정보가 노드에 저장되는 등 적절하게 처리되고 있다.

한편, 아크의 속성가운데 길이와 다각형의 속성가운데 둘레, 면적은 속성자료에 기록되어 있지 않아도 좌표를 이용하여 계산해 낼 수 있는 것들이다. 이들 자료를 자주 사용하는 경우에는 사용할 때마다 다시 계산한다는 것은 비경제적일 수 있으나, 자주 사용하지 않는 경우에는 용량의 측면에서 비경제적이다. 따라서, 프로젝트의 특성에 따라 설계자가 그 field의 설정을 결정해야 한다.



Nodes:

Number	Easting	Northing	Traffic Control	Crosswalk
1	126.5	578.2	Light	Yes
2	218.6	581.9	Sign	Yes
3	224.2	470.4	None	No
4	129.1	471.9	Sign	No

Arcs:

Number	From	To	Length	Pavement Condition	Lanes
I	4	1	106.3	Good	4
II	1	2	92.2	Fair	4
III	2	3	111.6	Fair	2
IV	3	4	95.1	Fair	2

Polygons:

Name	Owner	Arcs	Perimeter	Area	Zoning
A34	J. Smith	I,II,III,IV	405.2	10203	R-4
A35	R. White	III,VI,VII,XI	478.1	11562	R-4

〔그림 14〕 아크-노드 자료구조

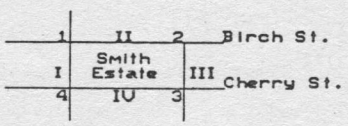
〔관계적 자료구조〕(relational structure)

관계적 자료구조는 아크-노드 구조와 매우 유사하다. 아크-노드 구조에서는 속성자료가 위상자료와 함께 저장되는 데 반해, 이 구조에서는 속성정보는 따로 관리된다.

〔그림 15〕에서 살펴보면 위상자료는 아크-노드 구조의 것과 동일하게 조직되어 있는 것을 알 수 있다. 속성정보는 테이블에 정리되어 있는데, 이 테이블을 릴레이션이라고 한다. 릴레이션의 행(rows)은 레코드(record) 또는 튜플(tuple)이라고 하고 열(column)은 필드 또는 속성이다.

위상자료와 테이블을 연결시키는 것은 두 자료가 공유하고 있는 필드(속성)로 그림에서는 Number와 Name이다. 이들 필드를 키(key)라고 한다.

속성자료를 독립된 화일로 관리하기 때문에 화일의 크기를 늘이거나 화일수를 증가시키면 속성자료를 쉽게 첨가시키고 관리할 수 있다. 특히, 일반목적의 데이타베이스 관리시스템이 관계적 자료구조를 지원하기 때문에 GIS에 연결시켜 사용하면 비용을 절감하고 융통성을 확장시킬 수 있다. dBASE Ⅳ, Oracle 등이 GIS에서 많이 이용되고 있다.



Nodes:

Number	Easting	Northing
1	126.5	578.2
2	218.6	581.9
3	224.2	470.4
4	129.1	471.9

Number	Traffic Control	Crosswalk
1	Light	Yes
2	Sign	Yes
3	None	No
4	Sign	No

Arcs:

Number	From	To
I	4	1
II	1	2
III	2	3
IV	3	4

Number	Length	Pavement Condition	Lanes
I	106.3	Good	4
II	92.2	Fair	4
III	111.6	Fair	2
IV	95.1	Fair	2

Polygons:

Name	Arcs
A34	I,II,III,IV
A35	III,VI,VII,XI.

Name	Owner	Perimeter	Area	Zoning
A34	J. Smith	405.2	10203	R-4
A35	R.White	478.1	11562	R-4

〔그림 15〕 관계적 구조

〔수치선분 그래프(DLG) 구조〕

미 내무성의 USGS는 수치지도자료 확보를 위해 많은 예산을 투자해 왔는데, 특히 이러한 자료의 국가표준을 개발해 오고 있다. 이가운데 수치선분 그래프(Digital Line Graph, DLG)로 알려진 표준 벡터 화일 포오맷이 있다. DLG의 자료원으로는 USGS 7.5 min(1:24,000) 및 15 min (1:62,500) 지형도가 사용되었다. 또한 1:2,000,000 축척의 지도계열을 기초로 DLG 화일이 제작되었다.

DLG화일의 자료내용은 서로 다른 주제 레이어로 구성되어 있다. 레이어1은 주경계 및 행정경계 등을 포함한 경계선, 레이어2는 하천·호수등 수자원, 레이어3은 교통망으로 구성되어 있다. 레이어4는 공공토지 조사 시스템(Public Land Survey System, PLSS)을 기초로 구성되어 있다.

DLG의 기본적 자료요소는 DIME, Arc-Node, 관계적 구조와 유사하다. 결절(node)은 선분의 종점, 또는 선분의 교차점을 나타내며, 선분의 특징 또는 형태를 표시하기 위해서 점(point)이 부가되기도 한다. 선분은 시점과 종점이 있고, 방향과 선분의 좌우에 있는 다각형(지역)을 지정할 수 있다.

또한 길이가 0으로 정의되는 특수한 degenerate line을 사용하여 내용이나 특징을 점으로 정의할 수 있다.

다각형은 선분으로 둘러싸이며 한 지점을 통하여 다각형의 특징을 나타낸다.

점·선·다각형이 위상관계와 위치를 나타내는데 사용되며, 각 요소의 속성정보를 부호화시키는 데에는 코드 시스템을 사용한다. 속성코드는 메이저 코드(major code)와 마이너 코드(minor code)로 조직되어 있다. 메이저

코드는 세자리 숫자로 구성되어 있는데 처음 두자리는 일반적 범주를 나타내고 세번째 자리는 부가적 특징을 나타낸다.

메이저 코드	범 주
020	기복 또는 고도
050	수 문
070	지표피복상태

마이너 코드는 네자리로 구성되는데, 첫자리는 0이다. 다음 세자리는 다음과 같다.

마이너 코드	범 주
001-099	결절(nodes)
101-199	다각형(areas)
201-299	선분(lines)
301-399	degenerate lines
401-499	일반 목적 코드(general-purpose)
601-699	설명 코드(descriptive code)

일반목적(general-purpose) 코드는 사상의 크기와 위치에 따라 점,선 또는 다각형으로 수치화될 수 있는 사상에 할당한 것이고 설명코드는 부가적 정보를 보충하기 위한 것이다.

DLG 자료화일에는 헤더 레코드가 처음 기록되는데, 헤더 레코드에는 화일 작성 일자, 투영법, 좌표 시스템, 화일내에 수록된 점·선·다각형의 갯수 등의 정보가 수록되어 있다. 결절의 자료레코드에는 위치등을 나타내는 설명

정보, 메이저 및 마이너 속성 코드, 문자(text string)이 포함되어 있고, 선분의 자료레코드에는 시점과 종점 좌우의 다각형 등과 같은 설명정보 (description), 선분을 나타내는 순차적인 x, y 좌표, 속성 코드, 그리고 문자가 수록되어 있다.

또한, 자료의 정확도에 관한 정보를 기록한 레코드가 비치되어 있는 경우도 있다. 벡터 자료구조의 개발과정을 고려할 때, DIME과 DLG는 같은 계열이며, 이들을 바탕으로 1990년 미국 센서스에 적용시킨 TIGER(Topologically Integrated Geographic Encoding and Referencing) 시스템이 개발되어 왔다.

2.4.5 벡터와 래스터 방법의 비교

래스터와 벡터방법은 지리정보의 모형화에서 뚜렷하게 구별되는 접근방법이다. 1970년대까지만 하더라도 두 방법은 서로·조화를 이룰 수 없는 대안이었으나, 최근 하드웨어와 소프트웨어의 발달로 상호변환이 쉬워졌고 하나의 시스템에서 두 방법을 동시에 수용할 수 있게 되었다. 이 두 방법의 장·단점을 비교하면 〈표1〉과 같다.

〈표1〉 **래스터와 벡터방법의 비교**

(벡터)

장 점

a) 현상적 자료구조를 잘 표현할 수 있고,

b) 자료구조가 축약되어 있으며,

c) 네트워크 연계로 위상관계를 잘 구축할 수 있고,

d) 그래픽의 정확도가 높고,

e) 위치와 속성의 검색, 갱신 그리고 일반화가 가능하다.

단 점

a) 자료구조가 복잡하고,

b) 지도중첩이 복잡하고,

c) 각 단위가 다른 위상형태를 갖기 때문에 모의(simulation)가 어렵고,

d) 표시기기(VDU)나 도화기가 비싸며,

e) 다각형 내의 공간분석이나 필터링은 불가능하다.

(래스터)

장 점

a) 자료구조가 간단하고,

b) 지도중첩이나 원격탐사자료와 연결시키기가 쉬우며,

c) 다양한 공간분석을 쉽게 할 수 있고,

d) 공간단위가 같은 크기와 형태를 갖기 때문에 모의가 쉽고,

e) 기술자체가 비싸지 않으며 발달속도가 빠르다.

단 점

a) 그래픽 자료의 양이 많으며,

b) 자료를 줄이기 위해 셀을 크게 하면 현상적으로 파악할 수 있는 구조나 정보를 잃어버릴 가능성이 크며,

c) 출력의 질이 나쁘며,

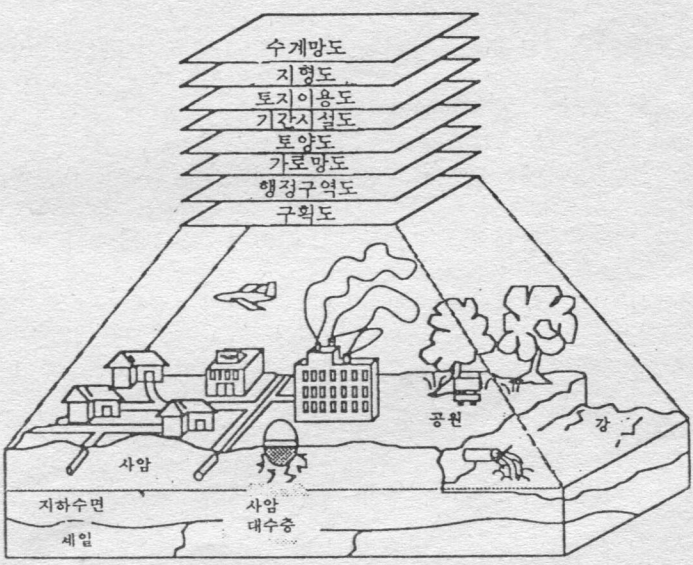
d) 네트워크 연계를 구축하기 어렵고,

e) 특수한 알고리즘이나 하드웨어를 사용하지 않으면 투영변환에 많은 시간이 소요된다.

2.5 GIS의 기능

GIS는 컴퓨터 기술을 이용하여 기타자료의 입력, 보관, 처리 및 출력을 보조하는 정보관리체계이다. 이러한 간단한 정의는 GIS를 단지 하나의 도구로 본 것인데, 만일 GIS를 하나의 시스템이 갖추어야 할 구조와 조직을 완비한 것으로 본다면 GIS는 컴퓨터의 하드웨어와 그것을 지원하는 소프트웨어를 운용하여 상호관련되는 기관과 전문인력을 완벽한 형태로 결합시키는 지리자료 처리 시스템이며, 또한 정보세계를 지원하는 계획을 세우는 일환이라고 볼 수 있다. GIS는 각종의 지리자료를 분산배치하고 표준적인 좌표체계를 통하여 하나의 완벽한 배열로 정리 및 결합시키기 때문에, 대상지역의 각종 자연 및 인문적 특성을 지리적 데이타베이스로 구축해 낼 수 있다〔그림 16〕.

일단 데이타 베이스가 구축되면, 사용자는 프로그램을 이용하여 분석과 시뮬레이션(Simulation)을 수행하게 된다. 이러한 작업의 결과로 얻은 합리적인 해답은 의사결정 과정의 보조자료로 활용한다.



〔그림 16〕 GIS와 레이어 구조

2.5.1 지리자료의 수치화

자료의 획득은 필요한 자료를 조사·확인하고 수집하는 과정으로 다양한 방법이 동원된다. 프로젝트에 필요한 자료와 자료원을 조사하고 조사방법도 정리한다.

야외조사 또는 방문·우편·설문조사 등을 통한 1차 자료가 이상적이지만, 비용과 시간이 많이들고, 때로는 불가능한 경우도 있기 때문에 2차자료의 사용이 불가피할 때가 많다. 지도·항공사진·인공위성 이메이지·타기관의 측량자료·간행자료 등이 많이 사용된다.

자료의 획득과정에서는 소요되는 시간과 경비의 계획이 철저해야 하며, 자료의 정확도·정밀도 등에 세심한 배려가 뒤따라야 한다.

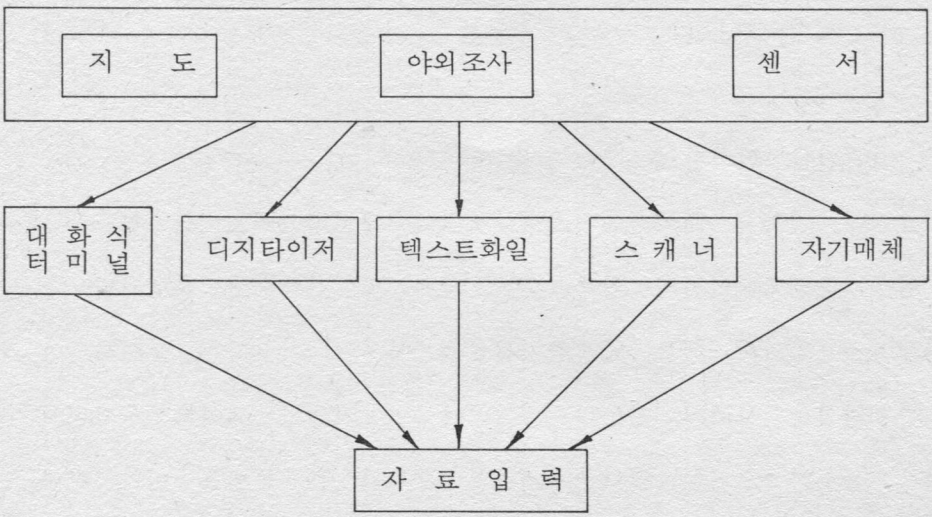
전처리는 GIS에 입력될 수 있는 형태로 자료를 처리하는 과정을 말하는데, 구체적으로는 데이터 포맷의 변환과 데이타베이스내의 위치지정을 가리킨다.

지도·항공사진·조사보고서 등으로 부터 정보를 추출하여 컴퓨터가 읽을 수 있도록 부호와(Encoding)하여 컴퓨터 데이타베이스에 기록하는 것 등이 데이타 포맷의 변환에 포함된다. 이 과정은 시간과 비용이 많이 들고 오류가 끼어들 수 있는 기회가 많다.

한편, 데이타베이스내에 오브젝트를 기록하고 위치를 지정할 때에는 일관된 체계에 따르도록 한다. 예를들면 레이어 시스템에서는 자료가 속한 레이어에 따라서 그 특성을 파악할 수 있기 때문이다. 또한 이 과정에 자료의 질을 통제할 수 있는 기준을 포함시키면 데이타베이스의 효용을 높일 수 있다.

지리자료의 수치화 작업은 기존의 처리 또는 도면자료와 그 속성자료를 변환시켜 컴퓨터가 받아들일 수 있는 형식으로 변환시키는 것을 가리킨다.

도형을 수치화시키는 일반적인 도구로는 디지타이저(Digitizer), 스캐너(Scanner), 비디오 카메라(Video Camera)등이 있으며, 지리공간이 가지는 고유한 속성은 키보드를 통해 입력시키는 경우가 대부분이지만 스캐너로도 속성자료를 입력시킬 수 있다〔그림 17〕.



〔그림 17〕 자료의 입력

수치자료화의 속도에 있어서는 비디오 카메라가 가장 빠르며, 그 다음이 스캐너와 디지타이저의 순이다. 반면에 가장 느린 것은 키-보드를 이용하는 경우이다. 자료의 질에 있어서는 일반적으로 스캐너가 가장 좋으며 그 다음이 디지타이저, 키-보드의 순이고 비디오 카메라가 가장 떨어지나, 자료의 종류나 이용목적, 그리고 경비와 소요시간 등을 고려하여 입력수단을 결정해야 한다.

이들 이외에도 원격탐사로 얻는 화상자료가 있는데, 이것은 지면의 상황을 이미 수치화시킨 것이기 때문에 직접 컴퓨터에 입력시킬 수 있다.

수치화하여 입력시킬 수 있는 주요한 지리정보로는 다음과 같은 것들이 있다.

a. 원격 탐사자료

b. 토지이용자료

c. 지적자료

d. 인구 및 주택자료

e. 지형자료(DEM, DTM)

f. 지질자료

g. 전신·전화자료

자료관리 기능은 데이타베이스의 제작과 데이타베이스에의 접근을 통제하며, 이 기능은 데이타의 입력, 갱신, 삭제 및 검색으로 수행된다. 최근 데이타베이스 관리시스템은 급격한 진보를 이루어 사용자는 오히려 이 기능에 큰 관심을 보이지 않는다. 그러나 이 기능이 제대로 수행되지 못할 때에는 사소한 오퍼레이터나 기계상의 오류가 큰 문제를 야기시킨다. 또한 이 기능에서 안전관리(Security management)도 이루어져야 하는데, 사용자의 수준에 따라 시스템과 데이타베이스의 접근할 수 있는 정도를 통제하는 프로시듀어가 장착되어야만 한다.

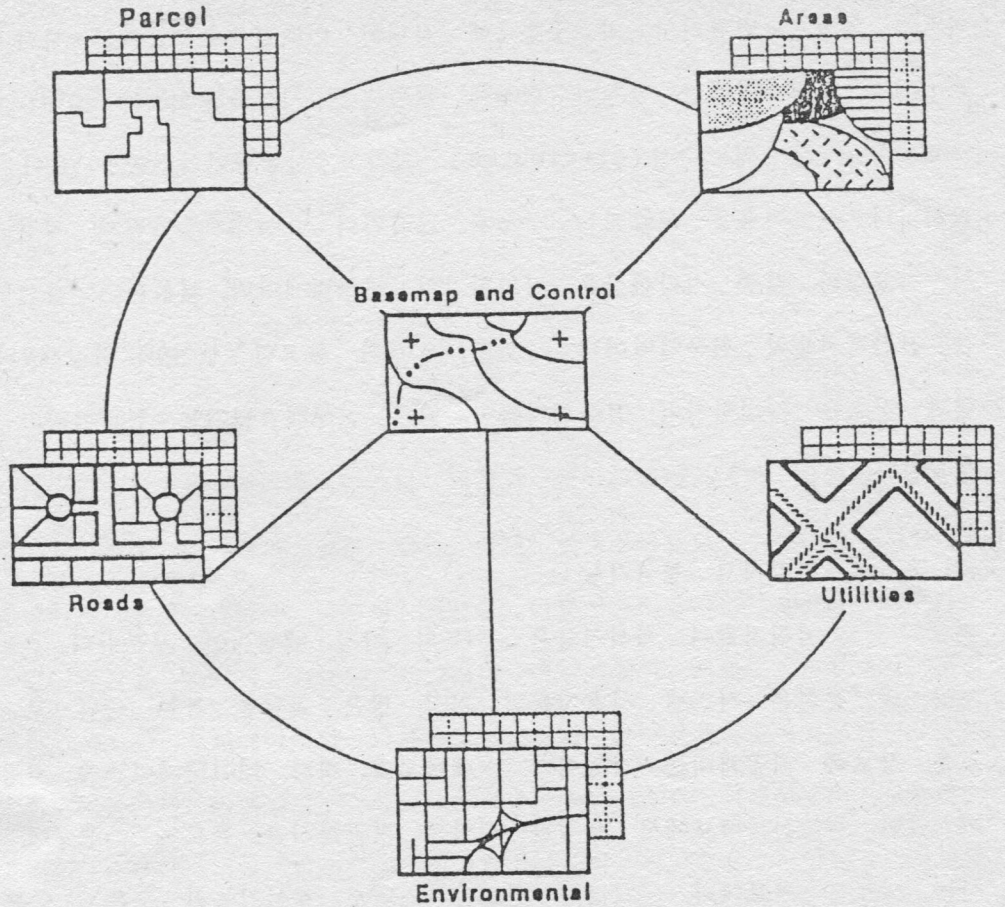
자료관리 시스템(DBMS)은 GIS에서 가장 중요한 부분으로 자료의 저장과 처리 및 분석의 전 과정의 효율성에 직접적인 영향을 미친다. 지리데이타베이스 시스템이란 지리정보의 기록체계를 말하는데, 이것은 지리정보를 효과적인 방식으로 저장하여 사용자에게 적시에 적절한 정보를 제공하기 위한 것이다.

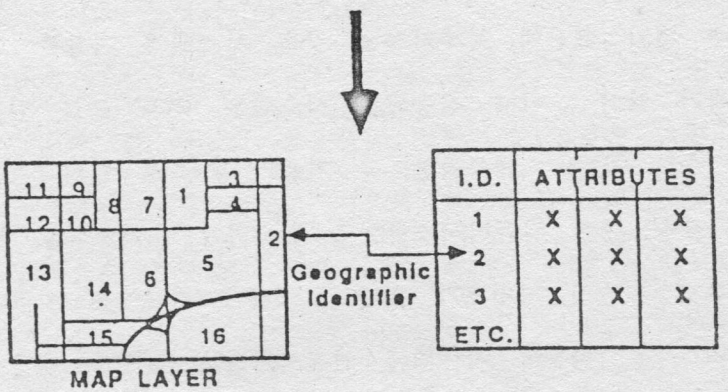
하나의 완벽한 지리적 데이타베이스 시스템이 되기 위해서는, 우선 지표상의 각종 지리현상의 특성이 개념적인 지리자료 모형(conceptual model of geographical data)속에 반드시 포함되고 이러한 개념적인 모형이 지리정보 시스템의 구조속에 포함되도록 구체화되어야 하며, 또한 지리현상

과 그 특성을 컴퓨터의 자료구조속에 적당히 안배할 수 있어야 한다.〔그림 18〕. 개념모형이란 특정한 현상에 관련되는 사상유형(geographic entity), 사상의 속성(attributes) 그리고 사상들간의 관계 등이 결합하거나 상호작용을 일으키는 구조를 가리킨다. 이러한 일련의 과정중에서 지표상의 각종 지리현상은 지리학자의 연구과제이며, 컴퓨터의 자료구조는 전산학자들의 전문영역이다. 또한, GIS의 중요연구분야인 지리자료의모형과 구조는 지리학자와 전산학자들이 함께 연구해 나가는 영역이라고 볼 수 있다.

2.5.2 지리자료 분석기능

지리자료의 분석기능은 GIS의 가장 핵심적인 부분이다. 또한 처리 및 분석은 시스템 사용자들이 가장 많은 관심을 쏟는 측면으로 새로운 정보를 가공해내는 기능이다. 예를들면, 지리 데이타베이스에 저장되어 있던 등고선 정보에서 특정한 지역의 평균경사를 유도하는 경우에 경사는 새로운 정보이다. 그러나 어느 시스템도 분석기능의 모든 종류를 구비할 수 없기 때문에, 특수기능을 갖춘 시스템에 자료를 전송해서 특수분석할 때가 있다. 이때 전송되는 시스템은 GIS 또는 일반적인 통계패키지일 수도 있다. 예를들면, 한 기관의 GIS로 부터 자료를 수치 분석기능을 갖춘 대형 컴퓨터시스템으로 전송하여 수치분석을 하고 다시 원래의 GIS로 전송하는 경우가 많다. 그러나 최근에는 이러한 특수기능을 모듈(module)로 만들어 시스템에 연결시키려는 전략을 채택하고 있는데, 이 전략으로 기존의 시스템을 기초로 분석기능을 점진적으로 확장시킬 수 있다. 일반적으로 geoprocessing이란 GIS의 처리 및 분석기능을 지칭한다. 분석기능을 다음과 같이 분류할 수 있다.





MAP LAYER AND TABULAR ATTRIBUTES LINKED BY GEOGRAPHIC IDENTIFIER

〔그림 18〕 지리자료의 개념모델

A. 공간자료처리 Spatial manipulation

a. 벡터방식과 래스터방식간의 변환 Vector/raster conversion

b. 좌표변환 Transformation

c. 좌표화 Coordination

d. 인근분석 Proximal or neighborhood analysis

e. 면적과 둘레의 계산 Area/perimeter calculation

B. 공간자료 분석 Spatial analysis

a. 다각형의 중첩 Polygon overlay

b. 다각형의 삭제 Polygon dissolve

c. 다각형 내부의 점 Point in polygon

d. 완충구역 및 회랑지역 설정 Buffering and Coridoring

e. 윈도우윙 Windowing

C. 수치지형 분석 Digital terrain analysis

a. 등고선 작성 Contouring

b. 경사와 주향계산 Slope/aspect

c. 하천유역 분석 Watershed analysis

d. 체적 계산 Volume calculation

e. 단면도 작성 Cross Section

f. 3차원 입체 화상 제작 3-D viewing

D. 네트워어크 분석 Network analysis

a. 적정노선 선택 Optimal path selection:route

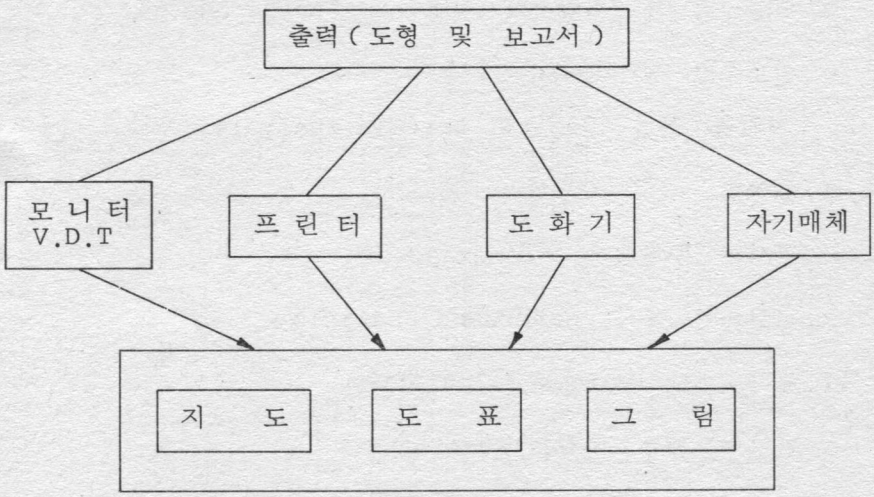
b. 유통량 시뮬레이션 Flow simulation:allocate

c. 시간 및 거리 구분 Time/distance districting:district

2.5.3 출력

출력에는 지역별 인구밀도·인구수와 같은 통계 테이블, 각종 주제도, 그래프 등이 있다. 또한 종이나 마이크로필름, microfiche와 같이 인식할 수 있는 형태와 자기테이프, 컴팩트 디스크등과 같이 컴퓨터가 읽을 수 있는 비가시적인 것이 있다. 또한 모니터에 일시적으로 나타나는 소프트카피(softcopy)와 종이나 필름에 기록한 하드카피(hardcopy)로 나누기도 한다〔그림 19〕.

최근에는 자료의 저장·교환에 국가적 또는 세계적 표준이 점차 적용되면서 자기매체로 출력하는 경향이 늘어가고 있다.



〔그림 19〕 자료의 출력

분석·처리된 지리정보를 사용자에게 정확히 전달하기 위해서는 표현수단(예, 지도)을 잘 설계해야 한다. 지도를 설계할 때는 반드시 지리정보가 표현하는 주제와 대상을 고려하여 정보의 양을 적당하게 조절하고, 사

용자가 적절한 정보를 편리하면서도 정확하게 판독할 수 있게 해야한다. 또한, 지리정보의 표현수단을 설계할 때는 출력시키는 도구의 성능을 고려하여 설계한 의도가 충분히 반영되도록 해야한다. 현재 일반적으로 사용되어지는 도구로는 프린터, 플로터, EPP (정전식 플롯 프린터), 모니터 등이 있다.

색상을 표현하는 데에는 모니터가 가장 아름답고, 그 다음이 Color EPP와 프린터이며, 플로터가 가장 뒤떨어진다. 한편 플롯(Plot)의 질에 있어서는 플로터와 모니터가 가장 뛰어나며, 그 다음이 EPP이며 프린터가 가장 뒤떨어진다. 선형의 벡터자료를 표현하기 위해서는 플로터나 모니터가 적절하며, 반면에 레스터식 자료를 표현하는 데는 프린터나 EPP를 사용하는 것이 좋다. 한편, EPP는 선형 벡터자료의 표현에도 이용할 수 있다.

정보표현(display)을 위한 설계에서 여러가지 도구들의 특성을 혼합시키면 양질의 출력을 사용자에게 제공할 수 있으며, 출력으로 나온 지리정보의 분석결과는 현상해설이라는 응용적인 측면의 기능도 발휘할 수 있다.

현재, GIS는 분산식 처리와 여러 네트워크를 결합시키는 방향으로 발전하고 있다. 이것은 각 부문별로 공정을 나누어 협업하는 방식과 마찬가지로, 각각의 네트워어크는 수치화된 지리자료를 개별적으로 관장하지만, 네트워어크간의 전송기능을 통해 상호 유통되어 결과적으로는 지리자료 자체가 공유되어진다. 컴퓨터 하드웨어나 GIS소프트웨어 역시 이러한 방향으로 발전해가고 있다. 이러한 분산식 처리방식의 활용을 통하여 수년 후에 GIS기술은 각 지방정부의 모든 업무에 보급 이용되어질 수 있을 것이다.

2.5.4 GIS의 개발

GIS의 개발은 일반적으로 타당성의 평가(feasibility assessment), 시스템의 설계 및 구입, 수행, 그리고 운용 및 관리의 4단계로 나눌 수 있다. 각 단계를 추진하는 데에는 많은 노력과 여러가지 활동이 요구된다. 각 단계의 활동은 다음과 같이 분류할 수 있다〔그림20〕.

1. 개념적 시스템 설계(Conceptual System design)
2. 시스템의 구입 및 설치(System acquistion and installation)
3. 데이타베이스의 개발(Database development)
4. 실무에 적용(Application development)
5. 훈련 및 업무수행(Training and service)

GIS개발의 5단계를 간략하게 살펴보기로 하자.

1단계 : 개념적 시스템 설계(Conceptual System Design)

GIS개발의 제1단계로서 GIS개발의 타당성과 시스템의 개념을 설계하고, 실행의 전반적인 계획을 수립하는 단계이다. 이 단계에서 GIS의 경영진, 전문가, 기술진에게 시스템의 목표를 설명하고(orientation), 필요사항을 수집하고 시스템 개발계획을 수립한다. 이 단계에는 다음의 6가지가 주요한 내용이다.

1. 연구계획 오리엔테이션(Project Orientation)
2. 필요사항 평가/타당성 조사(User Need Assessment/Feasibility Study)
3. 데이타, 시스템과 조직의 구비요건(Data, Systems and Organization Requirments)
4. GIS데이타베이스의 개념(GIS Database Concept)
5. 시스템의 개념(System Concept)

GIS Life Cycle	Developoment Tracks and Activities				
	Conceptual Design	System Acquisition and Installation	Data Base Development	Application Development	Training and On-Going Support
1. Feasibility Assessment	• Orientation • Needs Assessment • Requirements • Cost/Benefit				
2. System Design and Procurement	• Database Design • Application Definition • HW/SW Concept • Organization Concept	• HW/SW Specifications • Tender Documents • HW Vendor Contracting • HW/SW Acquisition	• Physical Database Design • Pilot Study	• Application Specification • Prototype Creation	
3. Implementation	• Implementation Strategy • Funding/Financing Strategy	• HW/SW Installation • System integration • System Programming	• Data Standardization • Data Conversion • Database/library Creation • Product Creation	• Application Creation • Documentation	• User Training • Administrator Training • Programmer Training • Application Training
4. Operations and Maintenance		• System Upgrade	• Update Procedure Definition • Service Bureau to Users	• Application Enhancement	• On-Going Education • Contracted Support • On-Site Support • Funding Program

〔그림 20〕 GIS의 개발

6. 실행 계획 (Implementation Plan)

2단계 : 시스템의 구입 및 설치 (System Acqusition)

제 2 단계는 하드웨어, 소프트웨어, 통신기기 등 시스템을 선택, 구입하고 설치한다. 연구기관에서 필요로 하는 연구계획에 맞는 시스템을 선정하고 설치하는데, 이에는 다음과 같은 실행절차가 있다.

1. 하드웨어의 선택과 구입 (Hardware Specification and Acqusition)
2. 하드웨어와 소프트웨어의 최종 결정 (Finalizing Hardware and Software Agreement)
3. 하드웨어 환경 조성 (Preparing Hardware Environment)
4. 하드웨어와 소프트웨어의 설치 (Installing Hardware and Software)
5. 적합성 평가 (Performing Acceptance Test)

3단계 : 데이타베이스의 개발 (Database Development)

제 3 단계는 지리정보 데이타베이스를 구축하는 단계이다. 데이타베이스를 완성하려면 오랜 시간과 많은 준비가 필요하다. 그러나 제 1 단계의 개념 설계를 통해 이는 대부분 해결되지만, 이 단계에서 구체적인 확인이 있어야 한다. 이 단계에서는 실험연구 이후에 완전 자동화를 실행한다. 이에는 다음의 8가지 실행단계가 있다.

1. 시험 설계 (Pilot Study Design)
2. 운용 데이타베이스의 구체적 설계 (Physical Design of the Operational Database)
3. 데이타베이스 점검 (Quality Assurance Program)
4. 자동화 시험 (Pilot Automation)
5. 자료의 수집 (Data Collection)

6. 자료의 편집과 오류제거(Data Compilation and Rectification)

7. 획득의 자동화(Data Automation)

8. 데이타베이스의 구축(Database Creation)

4단계 : 실무에 적용(Design and Development of User Applications)

이 단계에서는 GIS를 이용한 구체적인 실무적용을 계획한다. 이 활동은 GIS의 초기단계에 설정되어야 하지만, 적용될 때마다 여러번 검토되어야 한다. GIS개발의 초기단계에서는 이 단계의 활동을 제3단계와 동시에 수행한다.

1. 업무의 기능적 정의(Functional Definition of the Applications)

2. 업무의 설계(Application Design)

3. 업무의 구체화(Final Application Specification)

4. 업무의 개발(Application Development)

5단계 : 훈련 및 업무수행(Training and Ongoing Support Services)

이 단계에서는 GIS의 기술과 유지, 적용을 위해 사용자가 그 자원과 기술을 개발하는 단계이다. 이 단계는 제4단계와 병행하여 실행하지만 이것이 자리잡은 후에는 시스템의 지속적인 성장에 원동력을 제공하게 된다. 이에는 다음의 5가지 절차가 있다.

1. 훈련(Training)

2. 하드웨어와 소프트웨어의 보강(Augmentation of Hardware and Software)

3. GIS관리(GIS Management)

4. 지속적인 지원(On-site Support)

5. 재원 확보(Determine Financing Mechanisms)

2.6 GIS의 응용

막대한 분량의 정보를 관리하는 정부기관의 경우 소장자료의 80%-90%가 공간분포의 성질을 갖춘 지리자료이다. 따라서, 정부기관이 GIS의 최대 수요자이다. 그러나, 많은 개인기업이 업무의 기획과 추진과정에서 지리정보를 이용하여 문제를 해결하려는 경우가 적지 않다. 간단한 화물 운송노선의 계획에서부터 복잡한 과정을 거치는 공장부지의 선정이나 상점부지의 선정에 이르기까지 GIS는 뛰어난 문제해결 수단이 될 수 있다. 학술연구 부문에서는 공간정보와 긴밀한 상호관련성을 가지는 분야인 지리학, 지질학, 고고학, 토목공학, 수리학, 지역계획, 도시계획, 지적, 부동산, 그리고 조경건축 등에서 GIS를 이용함으로써 연구성과를 향상시킬 수 있다.

GIS이용의 예는 다음과 같다.

* 국토계획
  - 지역계획
  - 도시계획
* 토지 문제
  - 토지이용 계획
  - 토지자원 관리
  - 해안지역 계획
  - 지가 평가
* 자원분석 및 경영관리
  - 자연자원 분석 경영관리
    - 야생동물
    - 해양동물

· 삼림

— 광물자원 경영관리

· 석유

· 경제성 광물

* 환경관리

— 환경자원의 분석 및 관리

— 환경감시

* 수자원 관리

— 유역 관리

— 하천 유량

— 하천 범람치 계측

— 해일 및 파랑의 관측

* 교통운송도로망 계획

— 도로 노선의 선택

— 교통망 계획

— 교차로 계획

* 국방응용

— 군사기지 및 부지 선정

— 지세 분석

— 전술 모의 훈련

— 행군로 계획

— 직사 및 곡사화기 배치

* 사회안전분야

— 경찰 차량, 소방차, 구호 차량 등의 순시 및 이동노선 분석

- 경찰관서, 소방서 부지 선정 및 관할구역 계획
- 소화전 계획 관리
- 삼림 감시탑부지 선정
- 산불예방
- 자연재해 구조

* 공공시설 관리
  - 공원부지 선정 및 계획
  - 가스 상수도 전화 전기관로 배선
  - 가스 저장 및 가압소 등의 입지 선정
  - 학군계획

* 지 형
  - 지질 상황 모의
  - 토목량(절토 및 성토량) 계산
  - 경사도와 주향분석

* 교통 관리
  - 자동차 컴퓨터 지도
  - 최적 운행 노선 선택

* 상권 및 시장분석
  - 산업지구 분석
  - 산업 수요 및 공급

* 기 타
  - 고고학, 고증
  - 부동산 관리(vector data + videography)
  - 인구 및 취업시장 분석

3. GIS의 발전-회고와 전망

3.1 역 사

지도 이용자들은 오랫동안 대량의 지리정보를 수록하고 판독하는 방법에 있어서 많은 고심을 해왔다. 그런데, 1950년대에 컴퓨터의 발명으로 이러한 문제를 해결할 수 있게 되었으며, 지리정보는 정보화시대의 매우 중요한 분야가 되었다.

3.1.1 발전 배경

지리정보 시스템은 다음과 같은 부문의 발전을 통해 성장해 왔다.

a. 컴퓨터 하드웨어 기술의 진보

b. 수치화 도구의 발전과 지리공간자료의 내실화

c. 컴퓨터제도 분야의 발전과 가격하락

d. 데이타베이스 관리 기술의 발전

e. 수치화 및 공간분석 기술의 발전

3.1.2 성장과정

3.1.2.1 1960년대

컴퓨터 설비가 무척 고가였으며 도화(Plot)시스템의 성능에 한계가 있으면서도 가격은 매우 높아, 이 단계에서의 발전은 모두 공공기관에 의해 이루어졌다. 예를 들자면, 캐나다 정부의 GIS와 60년대 말의 미국의 호구조사 기본도 시스템(GBF/DIME FILE)등이 있다. 한편, 이 시기에는 학술계통에서는 대다수가 레스터식 자료분석을 위주로한 GIS를 발전시켰는데, 함부르크 대학의 GRID, 메릴랜드 대학의 MANS등을 꼽을 수

있다. 그후 캐나다의 GIS와 미국의 DIME은 모두 벡터식 자료를 처리할 수 있게 되었으며, 특히 DIME은 공간적 위상관계(Topology)를 수용하고 있었다.

3.1.2.2 1970년대

북미에서 환경문제는 중요한 관심의 대상이었고, 정부는 GIS와 같이 계획입안에 활용할 수 있는 지원도구를 긴급히 물색하였으며, 이를 통하여 환경정보를 파악하고 신속하고도 적합한 해결책을 구하려는 노력이 있었다. 이 단계에서는 컴퓨터와 그래픽 주변기기 시스템의 성능이 향상되고 한편으로는 가격이 하락하면서, GIS는 미국 연방정부나 큰 기관의 전유물에서 벗어나게 되었고 많은 유관기관에서도 GIS를 발전시킬 수 있게 되었다. 그리고, 지리정보연구회의 학술행사가 경험의 교류와 저변확대에 큰 기여를 하였다. 국제지리학회는 두 차례의 지리정보연구회를 개최하였고, 그후에는 Auto-Carto학술회가 국제적인 컴퓨터 지도학회와 GIS학회를 이끌어 오고 있다. 이 단계에서 GIS는 도약을 보였으나, 대부분 래스터 방식을 위주로 하였기 때문에 데이타베이스 시스템을 결여하고 있는 경우가 많았다.

당시의 응용분야는 크게 두 가지로 나누어지는데, 하나는 천연자원 관리 분야로 그 정밀도는 비교적 낮아도 무난하기 때문에 대부분이 래스터방식을 사용하였고, 다른 하나는 토지나 공공시설의 관리 분야로 비교적 높은 정밀도가 요구되기 때문에 점차 벡터 방식의 사용이 증가하게 되었다.

70년대 말에는 GIS사용자가 증가함에 따라, 전문적인 GIS 소프트웨어 회사가 출범하기 시작했고, 이들을 통해 GIS는 더욱 전문화되어 갔다.

3.1.2.3 1980년대

80년대에는 GIS의 공간자료가 위상(topology)의 측면과

데이타베이스 관리 시스템을 포함하게 되었다. 또한, 워크 스테이션(Work station)의 도입으로 공간자료와 함께 다양한 속성 자료를 동일 시스템 내에서 분석할 수 있게 됨에 따라 경비의 절감은 물론이고 우수한 분석 결과를 얻을 수 있게 되었다. 그 밖의 추세로는 개별적인 공간분석 소프트웨어를 모듈(module)로 결합할 수 있는 조합식 도구 박스(Tool Box)의 소프트웨어 생산이다. 개별적 목표의 데이타베이스 역시 점차로 다용도의 데이타베이스로 통합되고, GIS의 상품화와 다용도화가 급속이 이루어지고 있다.

3.2 발전 전망

이미 개발 소개된 GIS의 다용도화와 상품화 추세 이외에도, GIS의 기본적 기능과 응용의 측면에서 보다 많은 새로운 발전이 있을 것이다.

3.2.1 기본적 기능의 측면

* 입력—측량 및 수치화 도구의 발전은 물론이고, 원격탐사정보는 더욱 다양하고 정확해지며 정보제공 또한 신속해질 것이다.
* 관리—신속한 운용과 계산에 더욱 적합한 공간자료구조의 연구, 인공지능을 이용한 정보상담 방식의 개발, 반자동 정보관리 시스템으로 오차의 해결에 큰 진보를 기대할 수 있을 것이다.
* 분석응용—공간분석 능력의 증강과 아울러 모형화 기능이 강조되어 GIMS(Geographic Information Modelling System)로 이행될 것이다.
* 출력—더욱 다양한 형식을 통해 보다 효과적인 표현방식이 개발되어질 것이며, 자동지도제작 설계분야와 결합될 것이다.

3.2.2 응용측면

* 지리 데이타베이스—더욱 완벽하고, 신속한 운용을 지원해 줄 수 있을 것이다.

* 공간분석—real 3-D의 개발 등으로 응용분야가 더욱 넓어지고, 모형화가 강조될 것이다.

* 인식증가—보다 많은 사용자가 GIS가 제공하는 편리함과 효용성을 활용하고자 할 것이다.

4. 미 조사 통계국의 TIGER시스템

4.1 序

미 조사 통계국은 1981년 주요한 목표를 설정하였는데, 1990년 인구 및 주택 센서스에서 자료조사·통계표 작성(Tabulation)·조사결과 배포에 수반되는 모든 지도적·지리적 처리를 자동화시킨다는 것이었다. 이것은 센서스에 필요한 지도·주소참조화일(address range reference file)·지리사상간의 관계 화일(geographic relationship file)을 전혀 새로운 방법으로 제작하는 것이며, 이를 위해서 미국내의 모든 가로 및 도로, 도로명 및 가로 명칭, 주소, 철도, 강·호수 등과 같은 수자원사상(hydrographic feature), 경계선, 지명, 지역코드(geographic code)등을 전산화하여 하나의 데이타베이스를 통합하는 작업이었다. 그. 결과로 나타난 것이 GIS기술이 성취한 것 가운데 가장 성공적인 TIGER(Topologically Integrated Geographic Encoding and Referencing)화일이다. TIGER시스템 구축에는 약 3억 달러의 경비가 소요되었다.

전통적인 방법으로는 1990년 센서스를 감당할 수 없다는 판단으로 TIGER시스템을 제작하게 되었다. 다음은 그 이유를 정리한 것이다.

1. 사회가 다원화되고 정보사회로 진입하면서 대중의 정보요구는 더욱정교한 계수방법(enumeration methodology)을 요구하게 되었다. 이를 만족시키기 위해서는 조사단위를 더욱 세분하고 조사항목을 증가시켜야 한다.
2. 따라서 더 풍부하고 정확한 경계선 정보(boundary information)가 필요하다.
3. 이 정보는 조사구 요도에 적절하게 나타내야 한다. 30만의 계수관이

조사에 지참할 요도는 약 130만 도엽으로 도엽당 5～6장을 제작한다면 약 700만장의 요도가 짧은 기간에 제한된 장비와 인력을 통하여 마련되어야 한다.

4. 1970년부터 실시된 우편조사법(census-by-mail method)을 활성화시키기 위해서는 지리코드(geographic code)에 주소를 연결시켜야 한다.

4.2 TIGER/database file

데이타베이스의 개념적 설계에는 이용자의 요구 혹은 필요사항을 잘 살피고 그것이 적절하게 지원되도록 구조화시키는 단계가 필수적이다.

TIGER 데이타베이스는 자료수집과 통계작성에 적절하도록 지리적으로 구조화하고, 주소를 지리적으로 부호화(geocoding addresses)시키고, 통계국의 조사구요도 및 발간통계도(product map)의 자료원이 되도록 설계되었다. 이 데이타베이스는 개념적으로는 하나의 화일이지만, 군(county)별로 분할(partition)되어 있는 등 실제로는 TIGER county partition 화일, GEO-CAT (Geographic Catalog of Political/Statistical area)화일, national partition 화일, 그리고 잠정적인 워크 화일의 4가지 종류로 구성되어 있다.

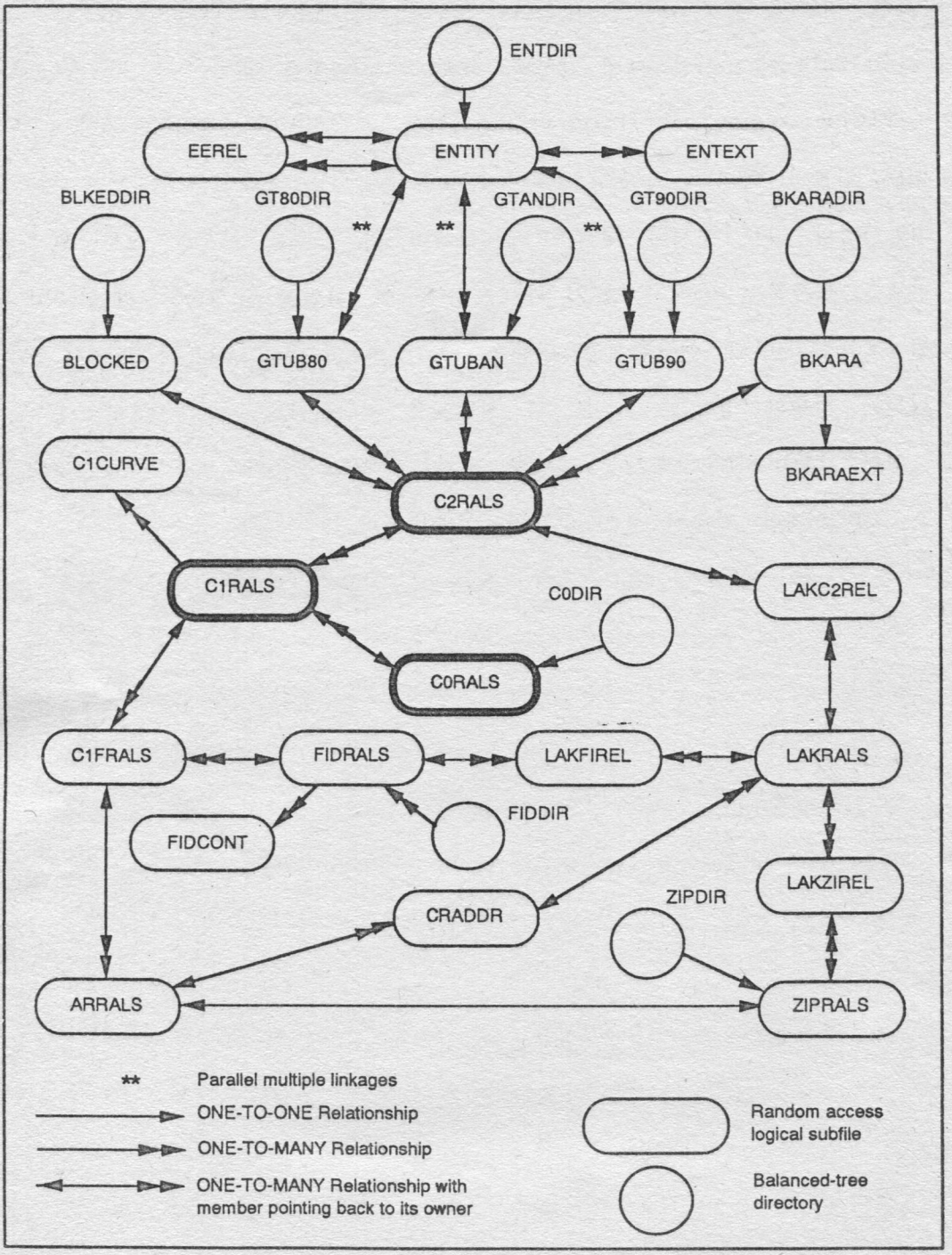
워크화일은 화일을 제작하고 갱신할 때 사용하는 일시적인 것이며, National partition 화일은 모든 군(county) 경계선의 좌표와 partition file identifier를 가지고 있어 참조화일의 역할을 한다.

GEO-CAT 화일에는 지리사상의 현재 및 과거의 정보, 명칭, 지리사상과 county partition file 사이의 관계 등이 수록되어 있고, 각 사상에 지리사상 코드(geographic entity code)를 부여하는 master referen-

ce의 역할과 함께 지리적 명칭이나 지리코드를 가지고 county partition file을 조사할 때에 rapid reference역할을 한다.

TIGER county partition file은 전체 TIGER 데이타베이스의 키(key)이며, 자료의 양이나 화일의 수에서 압도적이다. 이 화일에는 처리된 모든 지리사상의 좌표·코드와 TIGER 데이타베이스의 구조를 이루고 있는 위상관계를 수록하고 있다. 〔그림 21〕과 〈표2〉는 이 화일을 구성하고 있는 부화일(subfile)의 연결관계와 내용을 나타내고 있다. TIGER 시스템에서 출력되는 지도의 정보는 대부분 이 화일에서 지원된다.

한편, TIGER데이타베이스 화일의 자료원(data source)의 주요한 부분은 USGS에서 제공한 1:100,000의 DLG이다.



〔그림 21〕 하나의 **TIGER county patition** 화일내의 부화일

〈표 2〉 TIGER county partition subfiles

Name	Identification
ARRALS	Address Ranges
BKARA	1990 Block/Address Register Area (ARA)
BKARADIR	1990 Block/Address Register Area (ARA) Directory
BKARAEXT	1990 Block/Address Register Area (ARA) Extension
BLKEDDIR	1980 Block/Enumeration District Directory
BLOCKED	1980 Block/Enumeration District
C0DIR	0-Cell Directory
C0RALS	0-Cell (Zero-Dimensional Topological Cell)
C1CURVE	1-Cell Curvature
C1FIRALS	1-Cell/Feature Identifier Relationship
C1RALS	1-Cell (One-Dimensional Topological Cell)
C2RALS	2-Cell (Two-Dimensional Topological Cell)
CRADDR	Cross-Reference Addresses
EEREL	Entity-to-Entity Relationship
ENTDIR	Geographic Entities Directory
ENTEXT	Geographic Entities Extension
ENTITY	Geographic Entities
FIDCONT	Feature Identifier Continuation
FIDDIR	Feature Identifier Directory
FIDRALS	Feature Identifier
GT80DIR	1980 Geographic Tabulation Unit Base Directory
GT90DIR	1990 Geographic Tabulation Unit Base Directory
GTANDIR	Ancillary Geographic Tabulation Unit Base Directory
GTUB80	1980 Geographic Tabulation Unit Base
GTUB90	1990 Geographic Tabulation Unit Base
GTUBAN	1990 Ancillary Geographic Tabulation Unit Base
LAKRALS	Landmarks, Areas, Key Geographic Locations
LAKxxREL	Relationships between the LAKRALs and the C2RALS, ZIPRALS, or the feature FIDRALS
ZIPDIR	ZIP Code Directory
ZIPRALS	ZIP Code

4.3 Cartographic Extract

TIGER데이타베이스는 지리사상을 0-cell(point), 1-cell(line), 그리고 2-cell(area)로 분류하여 저장하고 있기 때문에 이 자료에서 지도를 출력하기 위해서는 지도오브젝트(cartographic object)로 조직해야 한다.

또한 하나의 지리사상을 지도 오브젝트로 정리하기 위해서는 여러단계의 처리를 거쳐야 하고 여러개의 화일을 참조해야 한다. 예를 들면 호수를 지도화하기 위해서는 8개의 화일에서 정보를 가져와야 한다.

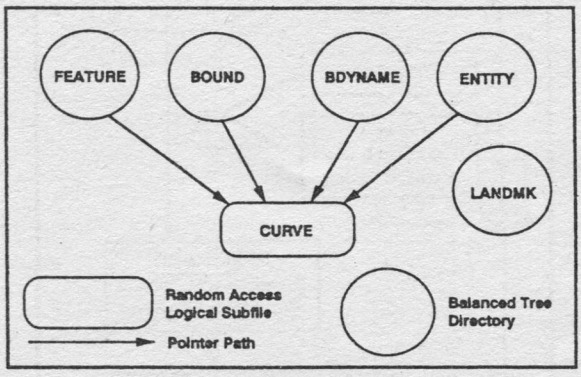
특히, 조사구 요도 제작에서 같은지역을 나타내지만 축척·내용·설계가 약간 다른 여러종류의 지도를 제작할 때 그때마다 이와같이 복잡한 절차를 거쳐야 하는 불편함이 문제가 된다.

이러한 장애를 제거하기 위해 제안된 것이 cartographic extract라고 하는 2차 데이타베이스다. Cartographic extract는 TIGER database에서 정보를 추출하여 구축한 것으로 〔그림22〕, 지도제작프로그램이 좌표스트링을 즉각 불러낼 수 있도록 구조화되어 있다.

Cartographic Extract를 이용하여 조사구 요도 제작은 완전 자동화되어 있고, batch processing으로 수행되고 있다. 지도의 축척, 지명 부착, 삽도(insert map)의 크기 및 축척, 도엽설계(map layout)등의 결정이 자동화되어 있다〔그림23〕.

예를 들면, 축척은 지역내의 지리사상 밀도(feature density)를 정량화하여 결정하는데, 조사구 요도의 축척은 1:5,000에서 1:70,000까지 500을 단위로 증감한다.

한편, 조사구 요도를 계산할 때에는 TIGER database의 수정을 거쳐서 cartographic extract를 갱신하도록 되어 있다.

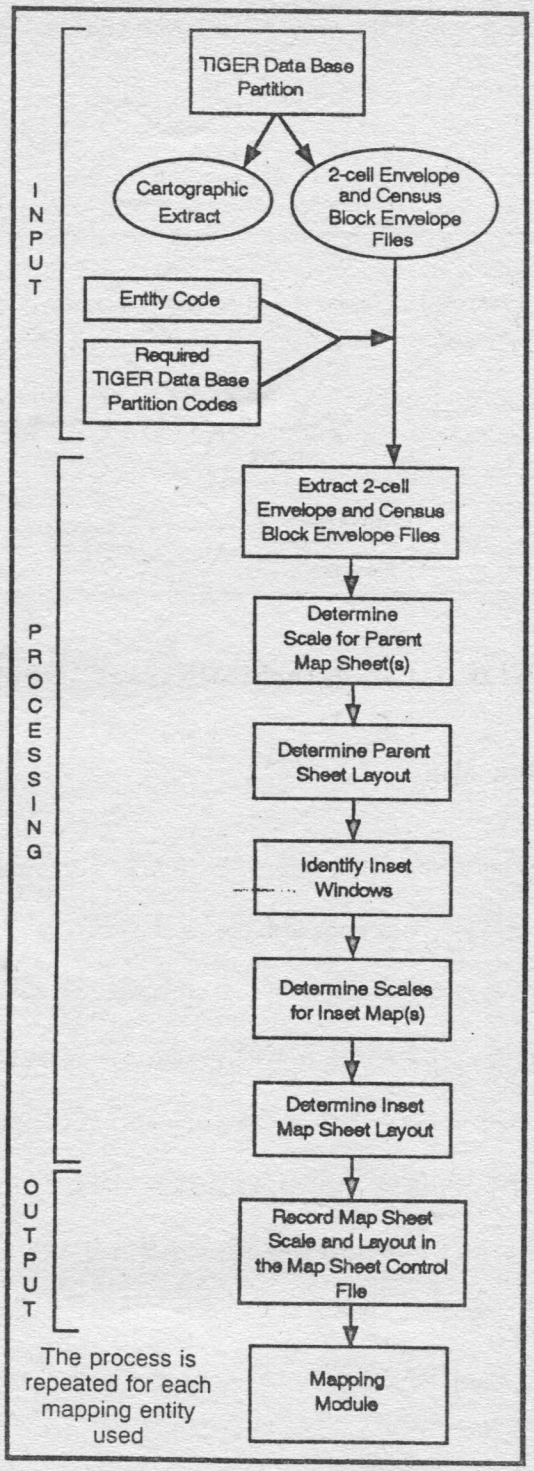


〔그림 22〕 **Cartographic Extract의 구조**

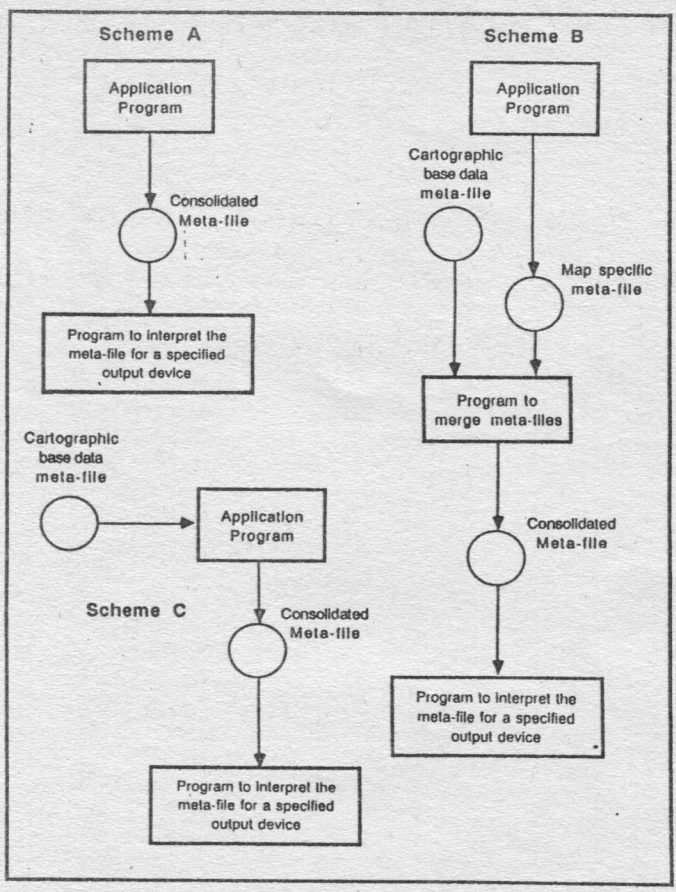
4.4 Publication Map

Cartographic extract가 조사구 요도를 제작하는데 필요하다면, 발간용 통계지도 작성에는 cartographic meta-file을 사용한다. 조사구 요도와는 달리 발간 통계지도는 종류가 작고 높은 질을 요구하기 때문에 interactive mode를 이용하며, non-interactive mode도 이용된다.

하나의 지도를 제작하기 위해서는 하나의 cartographic meta-file이 필요한데, 지도 제작에 필요한 meta-file에는 자료(object definition) 뿐만 아니라 function(subroutine)이 수록되어 있다. Cartographic extract와는 달리 축척·내용 등이 비슷한 다른 지도를 위해 제작된 meta-file을 가감하여 새로운 meta-file을 만들 수 있다.〔그림 24〕.



〔 그림 23 〕 축척, 도엽, 삽도결정의 자동화



〔그림 24〕 Meta-file과 통계지도의 제작

4.5 TIGER시스템의 성공 요건

지난 30년간의 GIS 역사에는 성공보다는 실패가 더 많았었다. 이가운데 TIGER 시스템은 GIS 기술축적이 가져온 개가라 할 수 있다. TIGER 시스템은 우리나라의 지역정보 전산화 또는 센서스 전산화 계획 및 수행에 많은 시사점을 주고 있다. 이 시스템을 성공으로 이끌었던 몇가지 요소를 정리하면 다음과 같다.

① Top management의 신기술에 대한 과감한 투자결단.

② GIS 기술의 성숙(hardware/software).

③ 조사통계국의 훌륭한 인적자원과 DIME file 제작을 통한 다년간의 축적된 경험.

④ 기존 직원(지도학자)의 경험과 신규채용 직원(프로그래머)의 신기술이 조화로운 결합을 이룸.

⑤ 국가 기본도가 준비되어 있었고, USGS가 DLG 등 수치지도 데이타를 제공하는 등 기관사이의 협조

〈부　　록〉

〔부록〕 미 조사통계국의 1990 센서스 조사용 지도

I. 1990 Census Field Maps

Field maps are designed to help with the collection of data for the 1990 decennial census. A map type may be used for a variety of decennial census field and office operations. To serve all requirements, map types often are modified to meet the needs of a specific operation. Field maps can be placed in three categories: 1) true field maps; 2) boundary collection/verification maps; and 3) early/limited use data review maps. True field maps are used only by Census Bureau staff for data-collection activities. Boundary collection/verification maps are used by Census Bureau staff and various local government and tribal officials for geographic entity boundary update and verification. Early/limited use data review maps are used by Census Bureau staff and various local government and tribal officials for additional geographic entity definition (census block equivalency) and for data review. These maps also are made available to the general public. The three categories of field maps are described below. (The highlighted numbers displayed parenthetically following each map type title show the number of map sheets being generated for the 1990 decennial census.)

A. True Field Maps

1. **Address Register Area (ARA) Map (311,835)**—A large-scale map produced for each ARA. Collectively, the groups for the four types of 1990 census enumeration cover the entire United States and its possessions: 1988 prelist; precanvass (TAR); 1989 prelist (update/leave); and list/enumerate. A limited number of geographic entity boundaries, including state, county, AI/ANA, and ARA, are shown. Areal features such as water bodies; linear features such as roads, railroads, pipelines, power transmission lines, and single-line hydrography; and landmark features such as mountain peaks, parks, and schools are displayed and named. Census block numbers are displayed. Listings identifying all streets and their associated ARAs within district offices may be used as a cross-reference to ARA Map sheets.
2. **Block Sketch Map (4,983)**—A large-scale map representing either a single census block or a cluster of census blocks. Features within the subject census block or census block cluster (including census block numbers) are identified. A list identifying all census block numbers in a specified number of surrounding tiers of blocks is included with each map for reference.
3. **Block Split Map (110,100)**—A large-scale map produced for each ARA. Cycle 1 maps include blocks split by geographic entity boundaries in selected TAR and 1988 prelist ARAs, while Cycle 2 maps include blocks split by such boundaries for all types of enumeration areas, except those ARAs previously mapped in Cycle 1 containing no new split blocks. In addition to the geographic entities shown on the ARA Map, other geographic entities shown on this version include MCD/CCD, incorporated place/CDP, UA, VTD, and CD. The boundaries of these tabulation areas, their names or identifying numbers, and tabulation block numbers are shown. Address range information is included in TAR areas only. Areal, linear, and landmark features are shown.
4. **County Locator Map (8,266)**—A medium-scale outline map used primarily by the DOs for planning, orientation, recruiting, and managing enumerator assignments. Map coverage includes all territory in each county. State, county, DO, AI/ANA, and ARA boundaries are shown. MCDs, incorporated places, and ANVSAs are identified by name only. All cartographic base features coincident with displayed boundaries are shown and major noncoincident features, such as roads and streams, are displayed. Cross-reference listings identifying the ARA numbers on a map sheet and the map sheets associated with a specific ARA number are provided with the map for reference.
5. **District Office Wall Map (4,827)**—Multiple full-size map sheets that can be mosaicked to display the service area of an entire DO. The Urban District Office Wall Map is a medium- to large-scale map produced for selected urban area DOs, while the Rural District Office Wall Map is a small-scale map designed for DOs covering a large geographic area, typically several counties. Within the DO, major features are shown along with geographic entities including county, CT/BNA, and ARA. Census block numbers are also shown on the Urban District Office Wall Maps.
6. **Postal Locator Map (1,404)**—A small-scale map produced for each mail out/mail back area located within an update/leave or list/enumerate county. State and county boundaries are shown. Selected cartographic base features are placed and named. The mail out/mail back area is highlighted by applying a screened shading over the adjacent fringe area.

B. Boundary Collection/Verification Maps

1. **Boundary and Annexation Survey Map (141,864)**—A map produced for each functioning county, MCD, and incorporated place. Maps are sent to local officials for update and certification of geographic entity boundaries. The 1988 BAS maps were designed to collect boundaries current as of January 1, 1988, whereas the 1990 BAS maps were designed to collect January 1, 1990, boundaries. Additionally, "BAS like" maps were produced for Puerto Rico, the Virgin Islands of the United States, and the Pacific Outlying Areas.
2. **Collection Insertion Map (74,063)**—Collection Insertion Maps were quad-based, fixed-scale maps created for editing insertion to the TIGER data base of 1990 census tracts and block numbering areas.
3. **Tribal Review Map (1,771)**—A medium-scale map produced for each AI/ANA and used by tribal officials for update and certification of boundaries.

C. Early/Limited Use Data Review Maps

1. **County Block Map (209,600)**—A medium-scale, county-based map, normally consisting of multiple sheets for each county. The maps portray the boundaries and names or identifying numbers of most geographic entities for which the Census Bureau tabulates data, as well as a wide range of cartographic base features. Census block numbers are shown.
2. **Governmental Unit Map (203,334)**—Governmental Unit Maps are produced for each MCD, incorporated place, and AI/ANA for which the Census Bureau tabulates data. They are medium-scale maps, identical in content and design to the County Block Map.
3. **TAR Geocoding Resolution/Work Place Spotting Map (62,123)**—A large-scale, census tract-based map used for researching unmatched addresses and identifying the locations of major employment sites. Geographic areas identified include state, county, AI/ANA, and CT/BNA. Areal features, linear features, and landmark features are shown. Address range information is included.

II. 1990 Census Data Product Maps

Data product maps are designed for use with the data the Census Bureau tabulates as a result of each census and sample survey. Most data product maps portray the boundaries of the political and statistical entities for which the Census Bureau tabulates data. Some also present the spatial distribution and relative magnitude of sets of census data. Data product maps can be placed in four categories: 1) data user maps; 2) printed reference outline maps; 3) printed thematic maps (page-size); and 4) printed thematic maps (GE-50 and GE-70 series). Both data user maps and printed reference outline maps are summary reference outline maps. Data user maps are electrostatically plotted, monochromatic, paper products. Printed reference outline maps are included in the published 1990 census reports and are printed in the traditional sense. Printed thematic maps are categorized according to size; page-size maps are approximately 8.5 x 11 inches while the GE-50 and GE-70 series maps are approximately 36 x 42 inches.

A. Data User Maps

1. **1990 County Block Map Series and Index (180,000)**—A comprehensive, multi-sheet, county-based reference map series. The maps portray and identify most geographic entities for which the Census Bureau tabulates data and a wide range of cartographic base features. Census block numbers are shown. An index to map sheets is provided for each county. The maps are prepared initially to accompany the special population summary prepared in accordance with P.L. 94-171. As such, they show the boundaries and identifying numbers of all voting districts submitted by participating state officials (90,000 map sheets). A second version of the map, on which voting district boundaries are omitted, is released to accompany the printed 1990 census reports and STFs (90,000 map sheets).
2. **Census Tract/Block Numbering Area Outline Map (4,200)**—A full-size (36 x 42 inches), county-based map series. Identifies MCD/CCDs, incorporated places/CDPs, AI/ANAs, and CT/BNAs. Only those cartographic base features coincident with CT/BNA boundaries are identified.
3. **County Subdivision Outline Map and Index (100)**—A state-based map series, generally available as a single, full-size map sheet. Identifies counties, MCD/CCDs, incorporated places/CDPs, and AI/ANAs.
4. **Voting District Outline Map (4,000)**—A full-size, county-based map series, produced for counties participating in Phase 2 of the 1990 Census Redistricting Data Program. Identifies MCD/CCDs, incorporated places/CDPs, and VTDs. Distinguishes between types of VTDs (actual and pseudo). All cartographic base features coincident with VTD boundaries are identified.
5. **Urbanized Area Outline Map, with selected base features (400)**—A full-size, UA-based map series with state, county, MCD/CCD, incorporated place/CDP, AI/ANA, and UA boundaries and names identified and major roads shown. All cartographic base features coincident with the UA boundary are symbolized and named. A fringe area is included and highlighted to distinguish the extent of the UA.

B. Printed Reference Outline Maps

1. **American Indian/Alaska Native Areas Outline Map (500)**—A page-size, AI/ANA-based map series. State, county, MCD/CCD, and incorporated place/CDP boundaries and names are shown in addition to the AI/ANA.
2. **American Indian/Alaska Native Areas of the United States (1)**—A two-page-size map of the United States showing state, county, and AI/ANA boundaries and names.
3. **Census Tract/Block Numbering Area Outline Map (4,200)**—A full-size (36 x 42 inches), county-based map series. The higher-quality version of the Census Tract/Block Numbering Area Outline Map available as a "data user" map.
4. **County Subdivision Outline Map and Index (300)**—A state-based map series, produced as a multi-sheet, page-size map series. Identifies counties, MCD/CCDs, incorporated places/CDPs, and AI/ANAs. An index to the map sheets is provided for each state.
5. **Congressional District Outline Map (550)**—A state-based, page-size map series bound in several Census Bureau publications and presented in the *Congressional District Atlas*. The boundary of each CD and the included counties are symbolized and identified. Selected incorporated places are shown. County-based inset maps are provided for counties containing two or more CDs. Cartographic base features coincident with CD boundaries are identified.
6. **Metropolitan Areas of the United States (1)**—A two-page-size map of the United States showing state and metropolitan area boundaries and names. Type style and boundary symbology distinguishes the three types of metropolitan areas—MSAs, CMSAs, and PMSAs.
7. **Public Use Micro-Data Sample (PUMS) Areas (18)**—Seventeen slightly overlapping sections covering the United States showing state, county, and PUMS "area" boundaries and names. An index to map sheets is produced.
8. **Regions and Divisions of the United States (1)**—A page-size map of the United States showing state, region, and division boundaries and names.
9. **State/County Outline Map (57)**—A page-size, state-based map series showing state and county boundaries and names. State capitals and selected places are included.
10. **State/Metropolitan Area Outline Map (57)**—A page-size, state-based map series showing state, county, and metropolitan area boundaries and names. The three types of metropolitan areas—MSAs, CMSAs, and PMSAs—are distinguished. State capitals and selected places are included.
11. **Urbanized Area Outline Map (400)**—A page-size, UA-based map series. State, county, MCD/CCD, incorporated place/CDP, and AI/ANA boundaries and names are shown in addition to the extent of the subject UA.
12. **Urbanized Areas of the United States (1)**—A two-page-size map of the United States showing state, county, and UA boundaries and names.

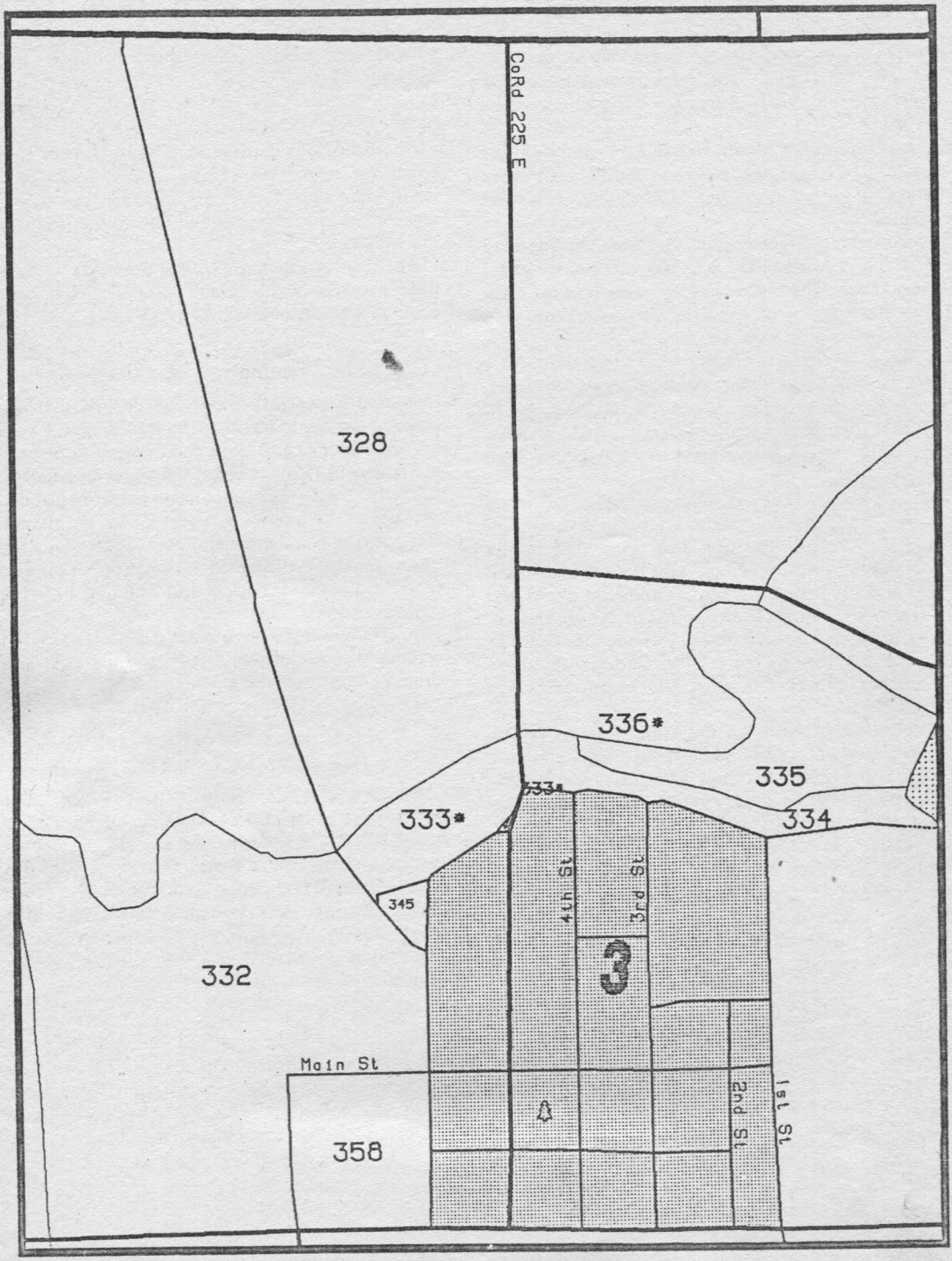
C. Printed Thematic Maps (Page-size)

1. **Assorted Titles (24)**—Page-size printed thematic maps of the United States. The two standard titles include Major Acquisitions and Dates of Admission of States and Centers of United States Population: 1790—1990. Additional choropleth maps are planned for each of the following report topics: population and housing counts; general population characteristics; general housing characteristics; social and economic characteristics; and detailed housing characteristics.
2. **Population and Housing Atlas (580)**—A proposed volume that would contain numerous page-size United States thematic maps.

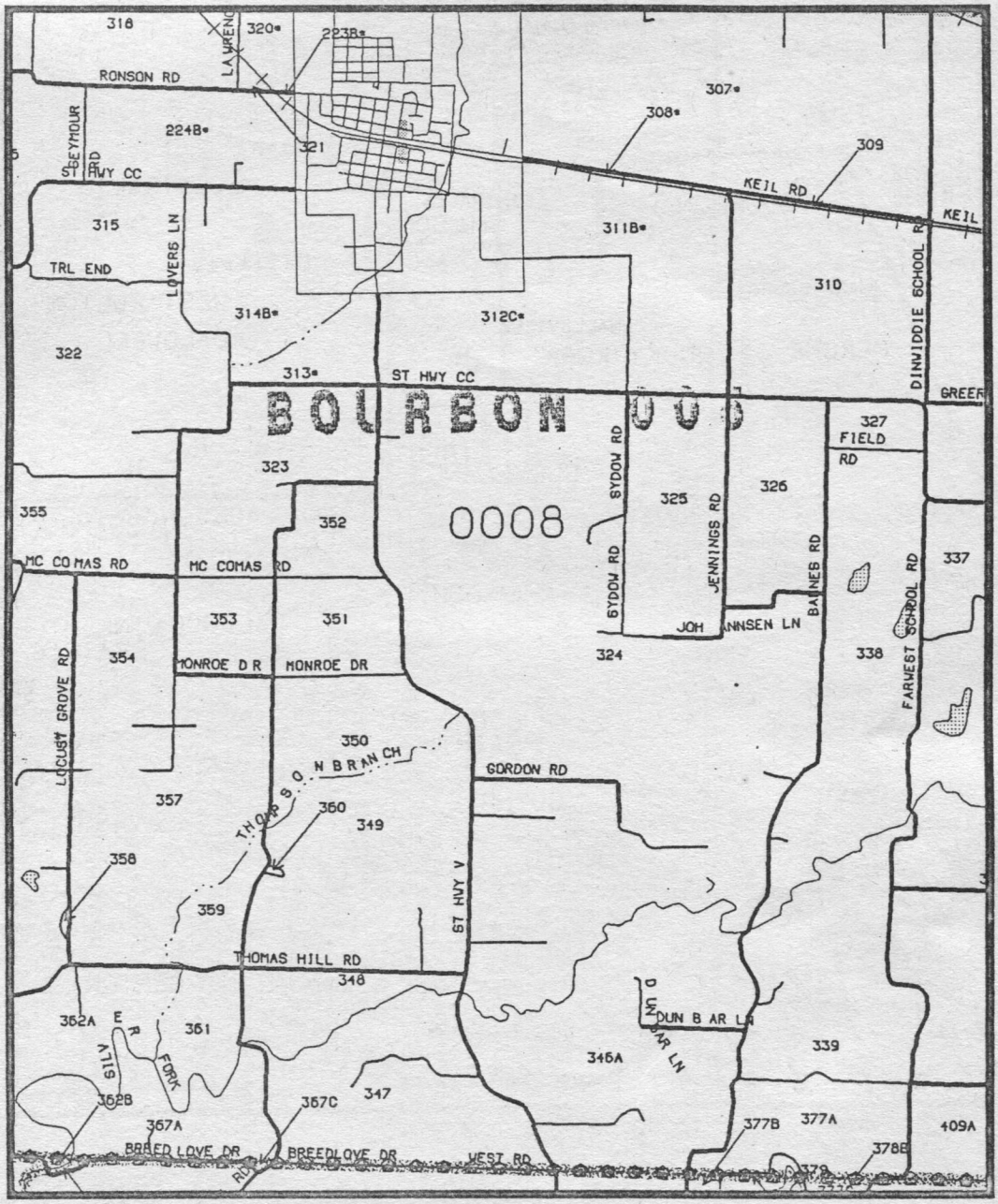
D. Printed Thematic Maps (GE-50 and GE-70 Series)

GE-50 and GE-70 Map Series (20)—Full-size, color maps of the United States, at a scale of 1:5,000,000 or 1:7,500,000. Standard titles include: Metropolitan Areas of the United States, Districts of the 103rd Congress, 1990 Population Distribution Map (Daytime View), and 1990 Population Distribution Map (Nighttime View). Additional choropleth titles are planned to depict selected 1990 census data by county.

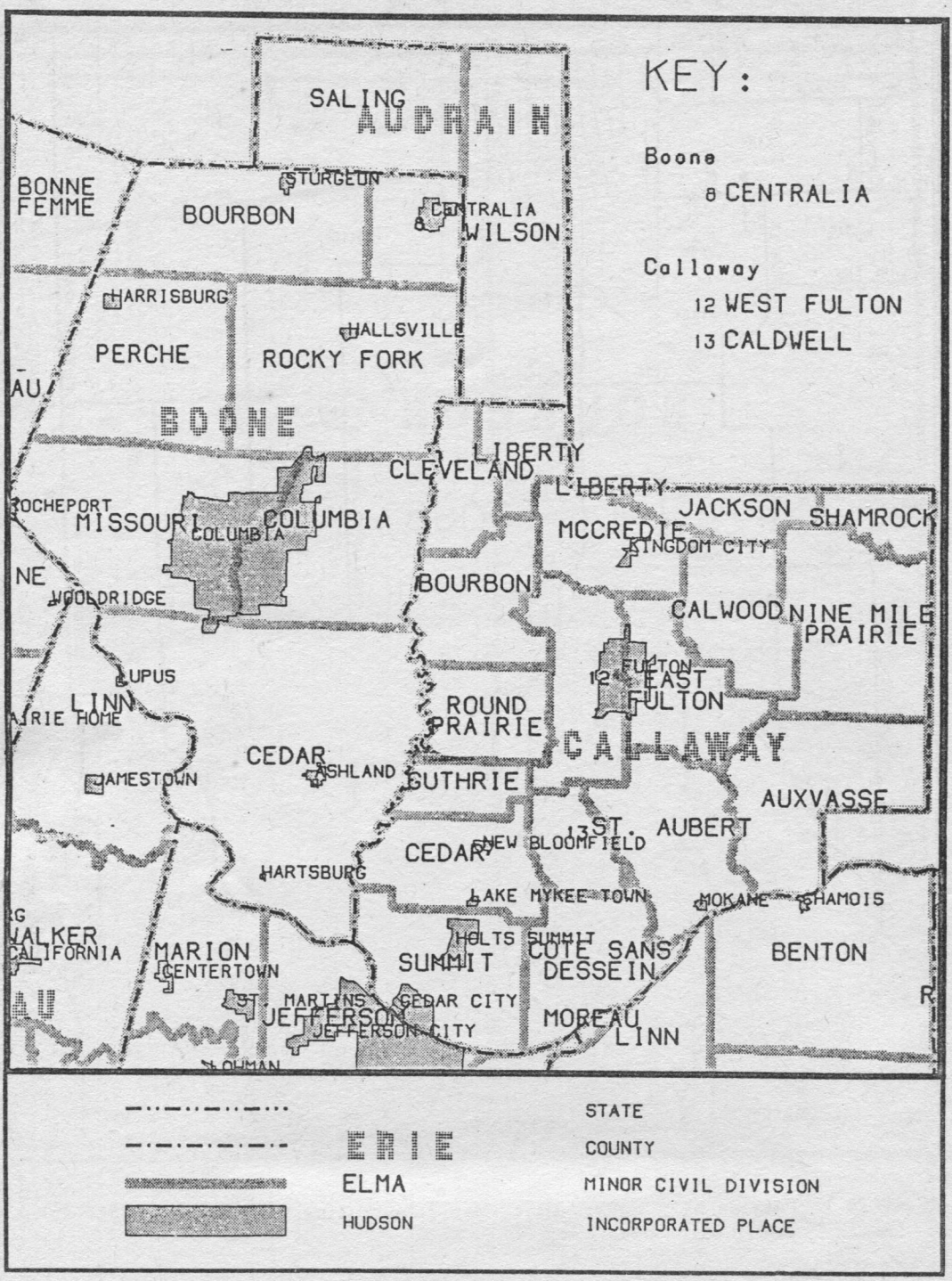
Additional information regarding the wide range of Census Bureau map products may be obtained by contacting Customer Services Branch, Data User Services Division, Bureau of the Census, Washington, DC 20233. Phone: 301/763-4100.



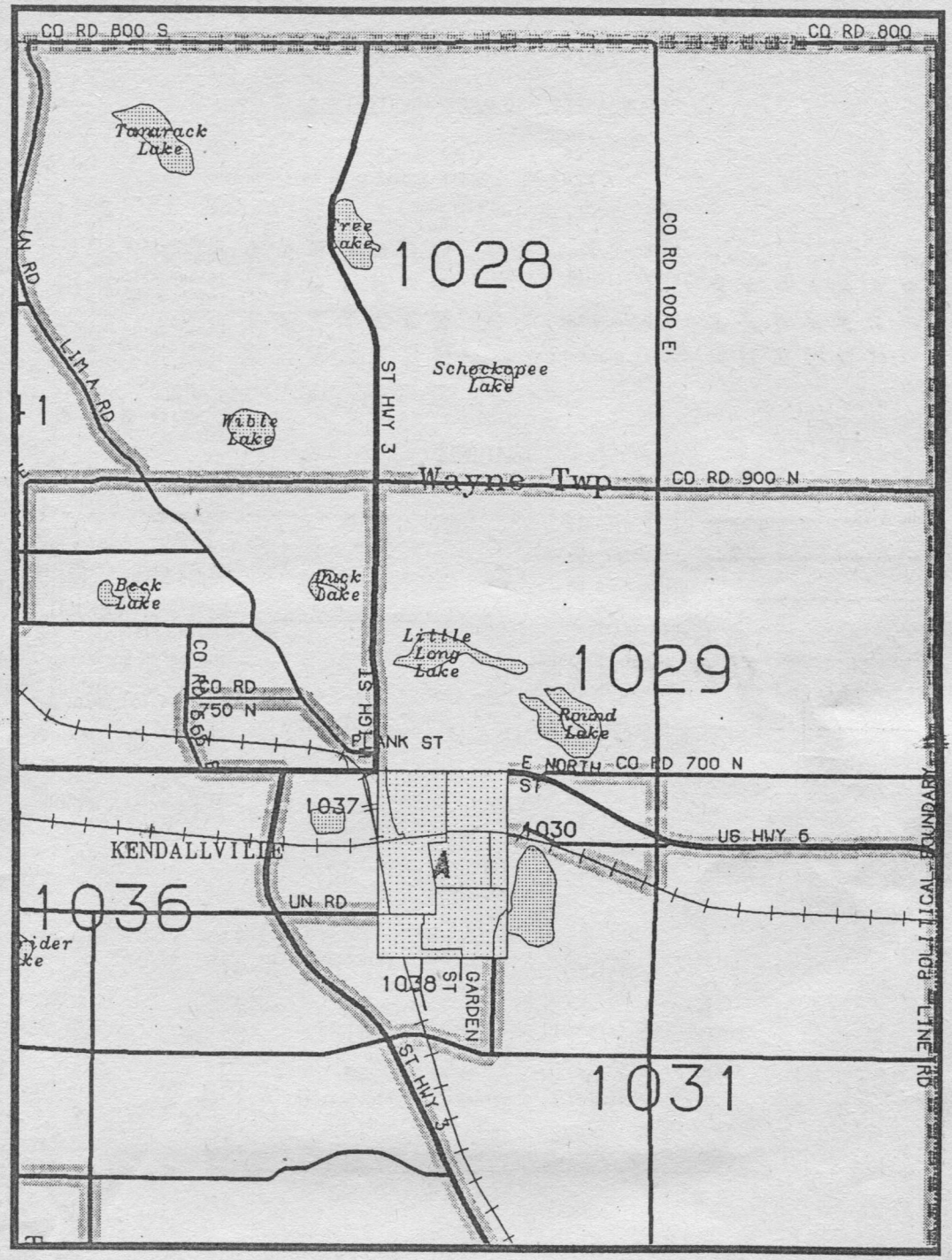
〔그림 25〕 **Portion of an Address Register Area illustrating an area for whieh the mapping software produced an inset by whole blocks.**



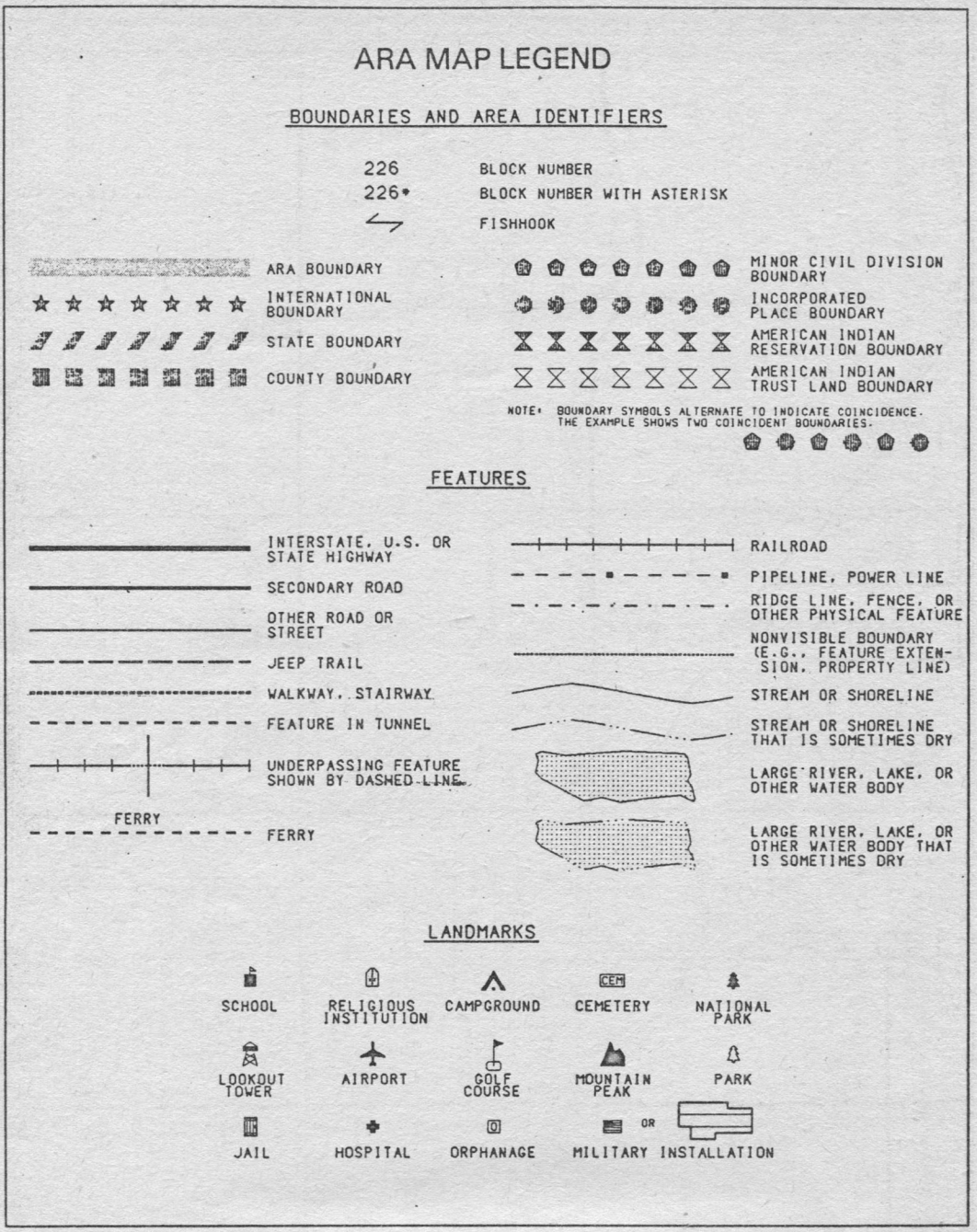
〔그림 26〕 **Portion of a County Block Map illustrating rectilinear insets for small-scale maps.**



〔그림 27〕 **Portion of a County Subdivision Map illustrating areal feature names Placement for governmental units and the use of key numbers.**



〔그림 28〕 **Portion of a County Locator Map illustrating the suppression of text within an inset area.**



〔그림 29〕 **Example of an Address Register Area Map legend.**

지역정보시스템 (GIS)

1990. 11

경제기획원 조사통계국

이 교재의 원고는 서울대학교 사회과학대학 지리학과 유 근 배 교수(지리학 박사)가 작성하여 제공한 것입니다.

목 차

1. 서 론

최근에 들어서, 지리정보 시스템(Geographic Information System, GIS)은 세계 각국의 정부기관, 개인 기업체, 정보산업계, 그리고 관련 학술분야 등에서 큰 관심을 보이는 분야가 되었다. GIS를 주제로 하는 각종 학술회의가 계속 열리고 있으며 그 반응 또한 대단하여, 매번 참석인원이 예상치를 초과하고 있다. 미국의 텍사스주 샌 안토니오에서 1988년 11월에 열린 "국제 GIS회의(GIS/LIS '88)"는 미국 원격탐사학회, 미국 지리학회, 도시 및 지역GIS학회 등 4개의 유력한 학술단체에 의해 공동으로 진행되었다. "세계에의 접근(Accessing the World)" 이라는 표어를 내세운 이 학술세미나에는 40여개국에서 3,000여명이 참석하였으며, 200여편의 논문이 발표되었다.

일본의 경우에는 1986년의 제1회 GIS학술세미나에 참석 예정인원이 100여명이었으나 세미나 당일 참석한 사람은 500여명에 이르렀기 때문에 세미나 장소를 더 넓은 곳으로 옮기지 않을 수 없었으며, 우리나라의 경우에도 금년 6월 서울대학교 국토문제연구소 주최로 열린 GIS워크샵에 참석한 인원은 예상을 훨씬 초과한 200여명에 달했다. GIS를 제작하고 있는 회사가운데 가장 유력해지고 있는 미국의 ESRI사에서는 10년전부터 자사 제품인 ARC/INFO의 USER CONFERENCE를 개최하고 있는데, 1981년 제1회때에는 16명이 참석하였으나 제10회였던 금년 5월에는 약 1,800여명이 모였다. 이제 세계 어느나라 할 것 없이 각국 정부기관, 정보산업계, 유관 학문분야등 사회 각 분야에서 GIS에 대해 관심을 가지는 사람의 수는 급격히 증가하고 있다.

그렇다면 GIS가 최근 몇 년 사이에 이와같이 각광을 받는 분야가 될 수 있었던 이유는 무엇인가? 우선 그 이유를 찾아보도록 하자.

전 세계적으로 급속한 산업화와 도시화를 겪으면서 지역간의 불균형, 도시인구의 폭증, 환경오염, 도시교통의 혼잡, 토지이용의 부적절 등 많은 문제를 안게 되었다. 또한, 정보사회로 진입하면서 정부기관을 비롯한 각종 기관은 정보공개에 대한 요구에 직면하게 되었다. 이러한 문제를 해결하기 위해서는 정확하고도 시기적절한 정보가 뒷받침되어야 하지만, 자료의 양이 방대해지고 종류가 다양해짐에 따라 전통적인 자료관리방법이나 기술로는 감당할 수 없게 되었다.

대부분의 국토정보는 현상의 장소·설명(속성)·시간으로 정의되는 지리정보(Geographic Information)로 구성되어 있다. 연구에 따르면, 미국 정부 각 부문의 정책결정과정에서 80%이상이 지리자료를 사용한다고 한다. 정부와 기업계의 입장에서 볼때, GIS를 통해 정부나 기업의 기구는 각각 담당하는 지역의 각종 현황을 정확히 파악함으로써 당면한 제반 문제들을 해결할 수 있고, 업무수행의 효율성을 제고시킬 수 있다.

또한, GIS는 급속히 발전하는 고도의 과학기술분야로서 유용성이 매우 높기 때문에 학술적으로도 급속한 진보를 보이고 있다. GIS는 지리학, 지질학, 자원공학, 임학, 해양학, 도시계획학, 지역계획학, 조경학, 토목공학, 경영학, 부동산 등과 같은 공간정보와 관련된 여러 학문분야에서 매우 효율적인 연구도구(Research tool)로 활용되고 있다.

한편, 정보산업체에서 볼 때, 그 시장규모가 연간 수억 달러에 달하며 매년 40%내외의 성장을 기록하는 GIS야말로 매력적인 분야가 아닐 수 없다. 따라서, 세계적인 컴퓨터회사들(IBM, DEC, PRIME, HP, SUN, APOLO, APPLE등)은 연구개발비용을 엄청나게 투입하면서까지 GIS시장에서 자사의 시장점유율을 확보하려고 노력하고 있다. GIS소프트웨어시장의 경우도 하드웨어와 마찬가지로 소프트웨어의 성능을 향상시켜 기술적인 측면

에서 우위를 점하려고 전력을 다하고 있다.

이와같은 정부나 기업체, 학계 그리고 정보산업체 전문인력의 움직임 속에서, GIS는 이 시대의 과학기술혁명에 있어서 심원한 획을 그을 수 있는 영향력있는 한 분야가 되었다.

2. GIS의 정의 및 자료

2.1 지리학과 정보시스템

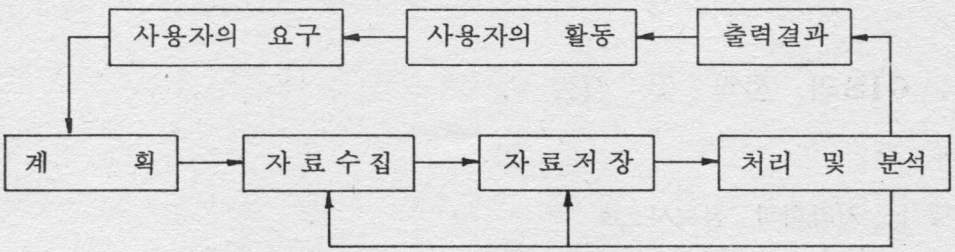
독일의 철학자이며 자연지리학자인 칸트는 학문을 크게세가지로 나누었다. 첫째, 특정한 대상, 군락, 또는 현상을 다루는 학문으로 생물학, 임학 등을 들었다. 두번째로, 시간의 경과에 따른 현상의 변화를 다루는 학문으로 역사학을 들었으며, 세번째로는 공간적 맥락(Spatial context)에서 자연현상 및 인문현상을 다루는 지리학을 거론했다. 전통적으로 지리학은 사람과 땅과의 관계에 중점을 두어왔고, 이 관계를 효율적으로 표현하는 지도가 key tool로 사용되어 왔다.

한편, 정보 시스템(Information System)이란 현상을 관측하고, 자료를 수집·저장·분석하여 의사결정에 유용한 정보로 압축하는 일련의 처리과정이라고 정의할 수 있으며, 자료원에서 정보의 사용까지를 다루는 end-to-end system이다〔그림 1 〕.

한편, 지도는 공간자료를 수집·분석하고 유용한 정보를 선별하여 지리좌표의 틀속에 정리하고 있으며, 정보를 명석 판명하게 전달하는 수단이기 때문에 정보시스템의 좋은 예라고 할 수 있다.

GIS는 공간좌표 또는 지리좌표(Spatial or Geographical Coordinate)에 참조된 자료와 이와 연관된 비공간자료를 다루는 것으로,"모든 형

태의 지리정보를 효율적으로 수집·저장·갱신·처리·분석 표시하기 위해 구축한 하드웨어 소프트웨어 지리자료 인적자원의 조직체"라고 정의할 수 있다. GIS는 서로 다른 축적(Scale)·시간(Time)·형태(Format)의공간자료를 연관시키고 통합시키는 수단이다.



〔그림1〕 정보 시스템의 개요

인류가 가지고 있는 지구에 대한 정보는 극히 한정되어 있고, 제한된 정보조차도 서로 긴밀하게 연관되어 있다고 볼 수 없다. 사람들은 현미경을 사용하여 분자수준의 현상을 관찰하기도 하고, 기상현상은 현상자체가 너무 크기 때문에 정지기상위성을 통하여 관측한다. 이 두가지 극단적인 축척(Scale)의 지구정보는 인간의 인식구조에서 얼마나 연관되어 있는가? 정보와 정보가 무한하게 연관되어 새로운 정보가 무한하게 창출된다면 지표공간은 어떻게 변모할 것인가? 공간정보를 처음으로 보편화시켰던 국세조사가 가져왔던 혁신을 생각한다면 그 내용과 범위를 상상할 수 있을 것이다.

GIS는 최근에 발달된 과학 기술 부문이기 때문에 아직 명확하게 정리되지 않은 부분이 많고, 처리·분석 및 출력 기능이 다양하고 적용되는 분야에 따라 강조되는 기능이 다르기 때문에 컴퓨터 지도학, 컴퓨터 응용 디자인(CAD), 원격 탐사 등의 시스템과 혼동하고 있는 사람들이 많다.

GIS와 컴퓨터 지도학은 많은 공통점이 있으나, 몇가지 점에서 큰 차이

를 보인다. 우선 공간자료의 변환에서 GIS에서는 위상관계(Topology), 공간적 위치, 비공간적 속성, 위치와 속성의 결합 등 폭넓은 기능을 가지고 있다. 또한 컴퓨터 지도학에서 비공간적 자료란 색, 선의 형태, 부호등을 가리키는데 반해, GIS에서는 토지이용의 종류, 소유권, 도로의 폭·차선·교통량, 식생의 종류 등을 말한다.

CAD에서는 수학적으로 정의되는 직선·곡선·도형으로 사상(Entity)을 표현하고, 사상의 Parent/Child의 관계를 이용하여 Object-oriented data base를 구축한다. 이에 반해 지리정보의 도형은 수학적으로 간명하게 정의될 수 없고, 자료의 양이 많고 복잡하여 Parent/Child의 관계로 조직될 수 없다.

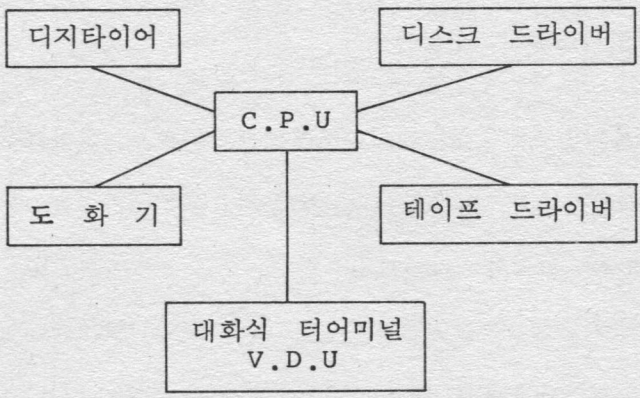
2.2 지리정보시스템의 구성요소

지리정보시스템은 하드웨어, 소프트웨어 모듈, 그리고 적절한 시스템 디자인의 세가지 요소로 이루어진다. GIS가 좋은 기능을 발휘하기 위해서는 세가지 요소가 균형을 이루어야 한다.

기능적 측면에서 GIS는 자료 획득(Data Acquisition), 전처리(Preprocessing), 자료관리(Data Management), 처리 및 분석(Manipulation & Analysis), 출력(Product Generation)의 다섯가지 기본기능으로 구성된다. 적용분야에 따라 강조되고 확장되는 요소가 다르지만, 이들 요소는 서로 연관된 것으로 연속적인 과정으로 파악해야 한다. 기능은 다음 장에서 좀 더 살펴보기로 하자.

지리정보 시스템의 일반적인 하드웨어 구성은 〔그림 2〕와 같다. 컴퓨터 혹은 CPU(Central Processing Unit)가 자료와 프로그램이 저장된 디스크 드라이브(Disk Drive)에 연결되어 있다. 디지타이저(Digitizer)

나 다른 주변 기기는 지도와 문서에서 추출된 자료를 숫자의 형태로 변환하여 컴퓨터에 보낸다. 도화기 (Plotter) 나 다른 출력기기는 자료처리 결과를 출력하는데 이용되며, 테이프 드라이브 (Tape Drive) 는 자료나 프로그램을 테이프에 저장하거나 다른 시스템과의 통신에 이용된다. 컴퓨터간의 통신은 특별한 자료선 (Data Line) 혹은 '모뎀 (Modem : Modulation —demodulation) '이라는 기기와 전화선으로 구성된 네트워어크 시스템 (Network system) 으로 이루어진다. 사용자는 컴퓨터와 주변기기들을 터미널의 화면 (Visual Display Unit : VDU) 을 통해 조작한다. 터미널에는 마이크로 컴퓨터 또는 지도를 신속하게 보여주는 특수한 하드웨어가 이용된다.



〔그림 2〕 지리정보 시스템의 주요 하드웨어 구성

2.3 GIS의 자료

GIS에서 처리하는 대상은 지리자료 (공간자료) 이다. 공간적인 분포의 특성을 가지는 것이라면 모두 지리자료라고 할 수 있으며, 지도상에 이러한 각종 지리자료를 기입하여 왔다. 흔히 사용되는 주제도로는 지형도, 지질도, 토양도, 지적도, 그리고 지하시설물 (상하수도, 지하 전화배선, 지하 가

스관 등)을 표시한 지도 등이 있다. 지리자료는 부호로 표시되는데, 그밖에 주소나 발생건수 등은 글과 숫자로 기재되기도 한다. 완벽한 지리자료가 되기 위해서는 공간자료(Spatial Data)와 속성자료(Descriptive or Attribute Data)가 함께 정리되어야 한다. 공간자료는 각종 지리현상의 위치와 형상 및 사상간의 공간상 상대적 위치관계를 말하는데, 지도상에서는 점, 선, 면을 사용하여 표시된다. 속성자료는 점, 선, 면으로 표시된 각종 지표공간의 특성을 말한다. 예를들어, 하나의 점으로 표시된 어떤 지표공간을 수질관측소라고 하면, 한쌍의(x , y)좌표로 그 위치가 공간자료로 변환한다. 속성자료에는 관측소의 명칭, 기호, 특정한 시각에 측정한 생화학적 산소 요구량(BOD)이나 용존 산소량(DO)등이 포함된다. 도로와 같은 선자료(線資料)는 일련의 점을 상호연관시켜 나타낸다. 속성자료인 도로의 명칭, 노폭, 노면재료, 교통량 등은 문자와 숫자로 데이타 베이스에 기재된다. 다각형 또는 면자료(面資料)인 행정구역도와 토양도 등은 하나 또는 여러개의 선분으로 구성되는 폐곡선으로 공간자료가 만들어지고, 속성자료로는 폐곡선으로 구분되어진 지역의 인구특성이나 토양의 성질 등이 기록된다.

지리자료 가운데 일부는 지형·토양·지질 등과 같이 그 속성이 쉽게 변하지 않는다. 그러나, 지적(地籍), 토지이용, 인구자료, 수질상황 등과 같은 자료는 수시로 변할 수 있는 것들이다. 따라서, 데이타 베이스의 내용을 어떻게 유지하고 갱신하는가 하는 것은 GIS의 구축과 발전에 있어서 반드시 고려해야 할 중요한 문제이다.

지리자료의 대부분은 정부기관이 수집·관리하는데 종류에 따라 수집과 관리하는 기관이 서로 다른 것이 일반적이다. 이와같이 행정 혹은 관료조직에 의존하는 지리자료의 수집 및 관리 방식은 관리의 측면에서는 매우 효율적이긴 하지만 타 기관에서 관리하는 지리자료의 활용이나 여러 기관에

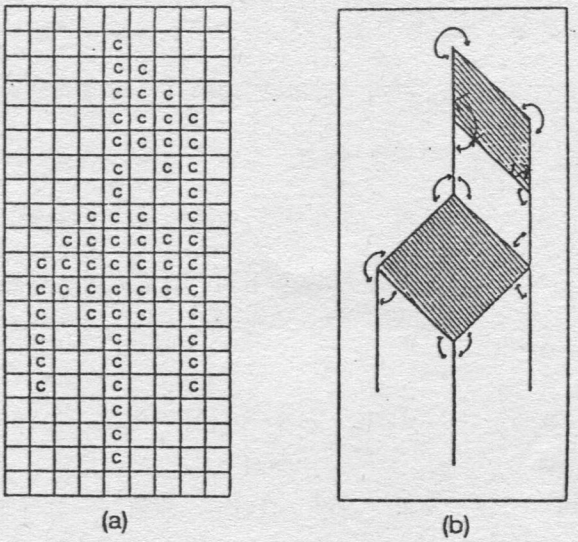
서 제각기 관리하는 여러 분야의 지리자료를 정리하고 통일시키는 데에는 어려움이 많다. 그 가운데 가장 심각한 것은 자료의 수집·분류·정리 등에 있어서의 표준이다. 이 문제는 동남아 각국이 모두 당면하고 있으며, 구미의 몇몇 선진국에서도 최근에서야 표준을 완성했다. 한편, GIS의 기술적인 발전은 이러한 문제의 해결을 용이하게 할 것이다.

2.4 GIS의 자료구조

사람의 눈은 형태를 인식하는데 매우 효과적이다. 그러나, 컴퓨터는 공간 패턴을 다루고 표시하는 방법을 정확히 지시받아야 한다. 공간정보를 표현하는 데에는 대조적이면서도 상호보완적인 두가지 방법, 즉 명시적(Explicit)방법과 내재적(Implicit)방법이 있다. 〔그림3〕은 컴퓨터에서 의자를 두가지 표현방법으로 나타낸 것이다. 명시적 표현이란 의자의 형태를 그리드(Grid)나 래스터(Raster)구조에서 점의 집합으로 나타낸 것을 뜻한다. 점의 집합으로써 책상이 아닌 의자를 표현하였으며 각 셀(Cell)은 'C'라는 코드값을 똑같이 가지고 있다. 실제로 'C'라는 것은 그 자체로 나타나지는 않고 수치나 색깔, 명암도 등으로 표현될 것이다. 의자에 관한 간단한 자료구조는 다음과 같이 볼 수 있다.

의자 ------ 기호화/채색화 ------ 셀X

내재적 표현은 시작점과 끝점, 그리고 연결의 형태로 표시된 선의 집합을 이용한다. 선의 시작점과 끝점은 의자의 형태를 표현하는 벡터(Vector)를 정의한다. 선과 선 사이의 포인터는 의자의 형태를 만들기 위해 선이 어떻게 연결되어 있는가를 컴퓨터에 알려준다. 자료구조는 다음과 같다.



〔그림 3〕 래스터(a)와 벡터구조(b)

〔그림3〕에 표현된 두가지 표현방법은 여러가지 점에서 차이를 보이고 있다. 첫째로 내재적 표현은 자료의 양이 적다. 즉 정보의 저장에 쓰이는 공간이 매우 적다. 벡터 표현은 11개의 좌표쌍과 14개의 연결포인터를 사용하는데 비해 래스터표현은 60개의 셀을 사용한다. 둘째로 벡터 표현은 외관상으로 래스터 표현보다 훨씬 산뜻하다. 〔그림 3〕의 경우 래스터 영상(a)으로 (b)와 같은 해상력을 가지려면 0.5㎜ 크기의 격자를 만들어야 하므로 이는 470개의 XY좌표가 필요하다. 세번째로 연결성의 정보(Connectivity Information)는 직접 공간검색에 이용될 수 있다. 그러나, 의자의 형태나 크기가 변하면 래스터식 표현은 벡터식보다 훨씬 빠르고 쉽게 고쳐질 수 있다. 래스터식 표현에서 정보의 수정은 단지 어떤 값을 지우고 새로운 값을 쓰기만 하면 된다. 벡터식 표현에서는 좌표값이

수정되어야 할 뿐 아니라 연결성도 다시 조직되어야 한다.

위상적 정보를 나타내는 가장 근본적인 이 두가지 방법은 다음과 같이 요약될 수 있다.

o 래스터 구조 : 좌표에 의해 정의된 셀 (Cell)의 집합. 각 셀은 속성값에 따라 독립적으로 기록된다.

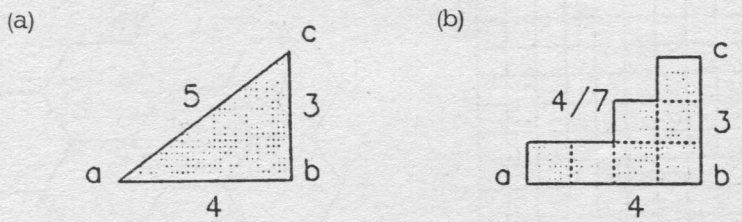
o 벡터 구조 : 점, 선, 면의 세가지가 있다. 점은 셀과 비슷하나 면적은 없다. 선과 면은 서로 연결되는 XY 좌표의 집합이며, 좌표는 주어진 속성과 관련시킬 수 있다.

2.4.1 래스터 자료구조

래스터 자료구조의 가장 간단한 형태는 그리드 (Grid), 셀 (Cell), 또는 픽셀 (Pixel)로 구성된 배열 (Array, Raster, Matrix or Lattice)이다. 각 셀은 행과 열의 값으로 참조되며, 지도화되는 속성의 값이나 유형을 나타내는 수치를 가지고 있다. 래스터 구조에서 점은 하나의 셀로 표현된다. 또한, 선은 어떤 방향으로 배열되어 인접하고 있는 셀들에 의해 구성 (표현)되며, 면은 사방으로 인접하고 있는 셀의 집합으로 표현된다. 이 자료구조는 지리정보를 표현한 2차원 자료공간을 연속적인 것이 아니고 분할된 (Quantized)된 것으로 가정한다. 따라서, 표현되는 사상에 비하여 셀의 크기가 클 때 길이와 면적의 계산에 큰 영향을 준다. 예를 들면 그림 4(a)에서 a와 c사이의 거리는 5이다. 반면 그림 4(b)에서 a와 c의 거리는 이것이 통과하는 셀의 변을 세는가 또는 셀을 세는가에 따라 7이나 4가 될 수 있다. 그림 4(a)의 면적은 6이며 그림 4(b)의 면적은 7이다. 셀의 크기에 따라서 정확도가 달라지기 때문에 수치화상처리 (Digital Image Processing)와 같은 분야에서

는 분할화된 표면을 연속적인 것으로 간주하여 미분 가능 함수를 적용하고 있다.

래스터 표현은 지리공간을 평평한 데카르트면으로 가정하며, 각 셀은 묵시적으로 정사각형의 지면을 나타낸다고 가정한다. 해상력 (Resolution)과 축적은 데이타베이스내에서 차지하는 셀의 크기와 셀이 실제 지표상에서 차지하는 면적의 관계에 따라 정해진다.



〔 그림 4 〕 래스터화가 면적과 거리에 미치는 영향

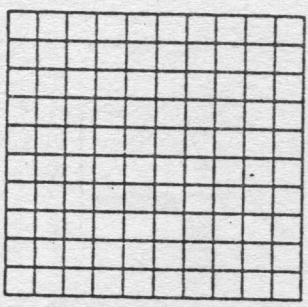
2.4.2 공간 분할 기법

평면을 같은 크기의 셀로 나누는 방법 (Tessellation) 에는 여러가지가 있다〔그림 5 〕. 가장 많이 사용되는 것으로는 정사각형이 있는데, 도트 매트릭스 프린터가 일반화되기 이전에는 라인프린터의 활자 형태에 따라 직사각형도 많이 사용되었다.

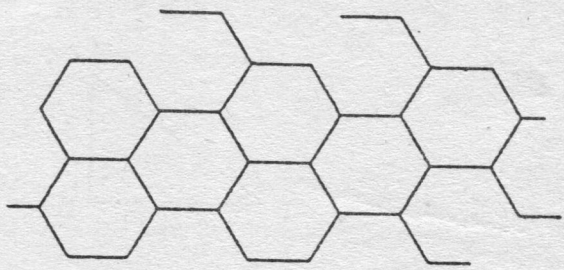
사각형 격자는 행렬로 간명하게 정리될 수 있다는 큰 장점이 있는 반면에, 인접격자에 대한 정의에 있어서 문제가 있다. 즉, 격자의 변에 인접한 4개의 셀을 인접격자로 보느냐 대각선 방향의 4격자도 포함시켜 8개로 보느냐 하는 것이다. 격자사이의 거리도 1과 1.41이 되기 때문에 또 다른 문제가 된다.

이에 대한 대안으로 정육각형법과 정삼각형법이 있다. 정육각형법은 인접격자의 정의가 명백하며 그 거리도 모두 같다는 장점을 갖고 있어 많은

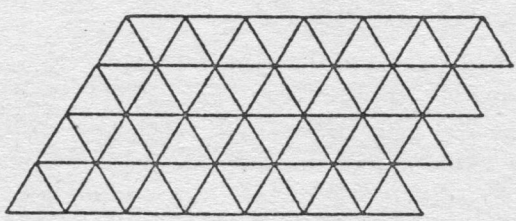
관심을 모으고 있다. 그러나 정사각형과는 달리 육각형은 같은 형태로 분할될 수 없기 때문에 해상력의 조정이 어려우며, 좌표화 또는 수치화하는 데에 부대공간 (Overhead)이 필요하다.



Rectangular



Hexagonal



Triangular

〔 그림 5 〕 공간분할기법 (**Tessellation**)

2.4.3 압축된 래스터 구조

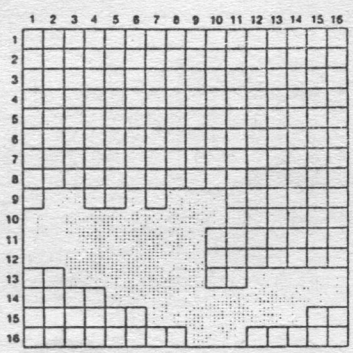
래스터 자료를 압축하는 주요방법으로는 체인 코드(chain code), 런 렝쓰 코드(run-length code), 블록 코드(block code), 그리고 사지수형(quadtree)의 네가지가 있다.

〔1〕 체인 코드

그림 6에서 경계선은 그 시작점과 동·서·남·북 방향(도엽의 위쪽을 북으로 했을 때)으로 순차 진행하는 단위 벡터에 의해 정의된다. 동은 0, 북은 1, 서는 2, 남은 3 등 숫자로 방향을 정의한다. 행 10, 열 1의 cell에서 시작한다면, 지역의 경계선은 시계방향으로 다음과 같이 입력된다.

$0, 1, 0^2, 3, 0^2, 1, 0, 3, 0, 1, 0^3, 3^2, 2, 3, 0^2, 1, 0^5, 3^2, 2^2, 3, 2^3, 3, 2^3, 1, 2^2, 1, 2^2, 1, 2^2, 1, 2^2, 1^3$

여기서 지수는 각 방향으로 중복되는 단계(cell)의 수를 표현하고 있다. 체인 코드는 압축시켜 저장하는 데에 매우 효과적인 방법일 뿐만 아니라, 면적과 둘레의 계산, 첨점(sharp turn), 오목성(concavity)등의 연산을 쉽게 할 수 있다. 반면에 합집합(union)이나 교집합(intersection)과 같은 연산은 어려우며, 또한 지역의 모든 경계선이 저장되어야 하기 때문에 자료의 중복이 불가피하다는 단점을 가진다.



〔그림 6〕 래스터 지도와 지역(region)

〔2〕 런 렝쓰 코드

각 지도단위에 속해 있는 점들을 각 열마다 왼쪽에서 오른쪽 방향으로 시작 셀과 끝 셀을 저장하는 방법이다. 〔그림 6〕은 다음과 같이 변환된다.

Row 9 2,3 6,6 8,10

Row 10 1,10

Row 11 1,9

Row 12 1,9

Row 13 3,9 12,6

Row 14 5,16

Row 15 7,14

Row 16 9,11

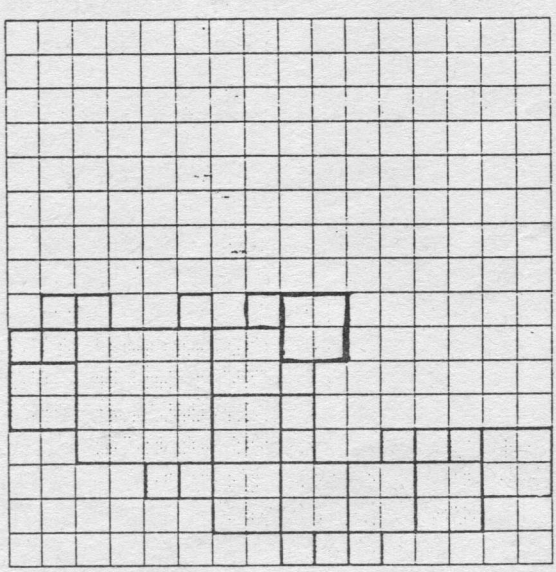
이 예에서 〔그림 6〕의 69(16×16)개의 셀이 정보의 손실없이 22개로 압축되어 입력된다. 따라서 자료를 저장하는 공간이 절약된다. 확실히 런 렝쓰 코드는 일대다수의 관계(한 지도단위가 다수의 셀로 구성된)가 있을 때에 자료저장이 용이하다. 특히 이것은 자료의 저장에 제한을 받는 소형컴퓨터에 적당하다. 그러나 너무 지나친 자료의 압축은 지도제작이나 자료조작에서 처리시간을 길게 만들 수도 있다.

〔3〕 블록 코드

런 렝쓰 코드의 착상은 지도화 되는 영역을 타일(tile)형태의 정사각형 블록을 사용함으로써 2차원으로 확대될 수 있다. 〔그림 7〕은 〔그림 6〕의 래스터 지도를 어떻게 정방형으로 표현하는가를 보여 준다. 자료구조는 세가지 숫자 즉, 원점(중간점이나 최하점)의 XY좌표와 사각형의 지름으로 되어 있다. 이것을 중간 축 변환(MAT:medial axis transform-

ation)이라고 한다. 〔그림 6 〕의 지역은 1개의 16단위 셀로 이루어진 정사각형과 9개의 4단위 정사각형, 그리고 17개의 1단위 정사각형으로 저장된다. 각 정사각형에 2개의 좌표가 필요하므로 이 지역은 57개의 숫자로 저장될 수 있다. 즉, 54개는 좌표값, 3개는 셀의 크기 또는 지름을 표시한다. 주어진 지역에서 맞춰지는 정사각형이 클수록 경계선은 더욱 단순해지기 때문에 이 방법은 효율적이다. 런 렝쓰 코드와 블록 코드 모두 크고 단순한 모양에는 효과적이나, 기본적 셀보다 불과 몇배의 크기를 갖는 지도단위로 복잡하게 구성된 경우에는 비효율적이다. 중간축 변환(MAT)은 합집합이나 교집합 연산을 하거나 지도 단위의 길이와 같은 특성을 포착하는데 장점이 있다.

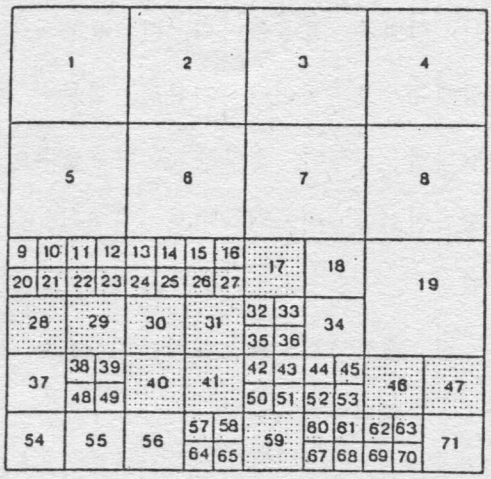
한편, 블록 코드나 런 렝쓰 코드로 저장된 자료는 몇몇 연산을 위해서 단순한 래스터 형태로 전환되어야 한다.



〔그림 7 〕 중간축 변환과 블록

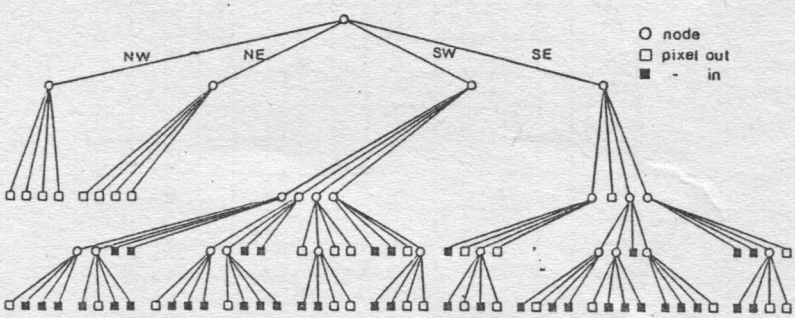
〔4〕 사지 수형 (quadtree)

자료를 축약하는 4번째 방법은 위계적으로 배열을 4분원 (quadrant)으로 나누는 것이다. 배열(지역)을 한 단계씩 밟아 사분원으로 분할시키면서 어떤 단계의 분원이 그 지역(mapping unit)에 완전히 포함되는가를 찾는다. 분할에서 가장 적은 단위는 하나의 화소이다.



〔그림 8〕 4지수형과 지역(region)

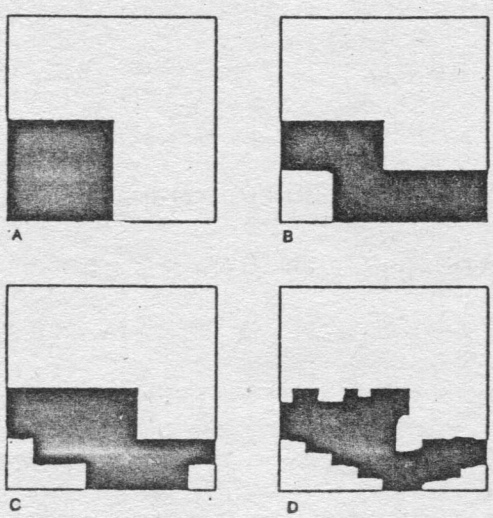
〔그림8〕에서는 〔그림6〕의 단위지역을 점진적으로 나누어 사분원 블록으로 채우고 있다. 이 블록 구조는 사지수형(quadtree)이라는 4개의 가지(degree)를 가지는 나무의 형태로 나타낼 수 있다.



〔그림 9〕 4지수형도

〔그림6〕의 4지수형은 〔그림9〕와 같다. $2^n \times 2^n$점의 배열은 4지수의 종결절(root node)이며, n의 높이(level)를 갖는다. 각 결절은 4개의 가지인 북서, 북동, 남서, 남동의 네개 분원을 갖는다. 엽결절(leaf node)은 더이상 분할되지 않은 4분원을 가리킨다.

4지수형은 다른 방법에 비해 많은 장점을 가지고 있다. 지역의 표준성(standard region property)이 효율적으로 계산된다. 4지수형은 가변적인 해상력을 가지는 배열(variable resolution array)이기 때문에 정보가 비어있는 부분을 단순하게 처리하면서도 정보가 있는 부분(면적)은 자세하게 다룰 수 있어 용량을 절약할 수 있고 검사시간을 단축시킬 수 있다〔그림 10 (a-d)〕.



〔그림 10〕 4지수형구조가 각 수준(n)에서 나타내는 지역형태

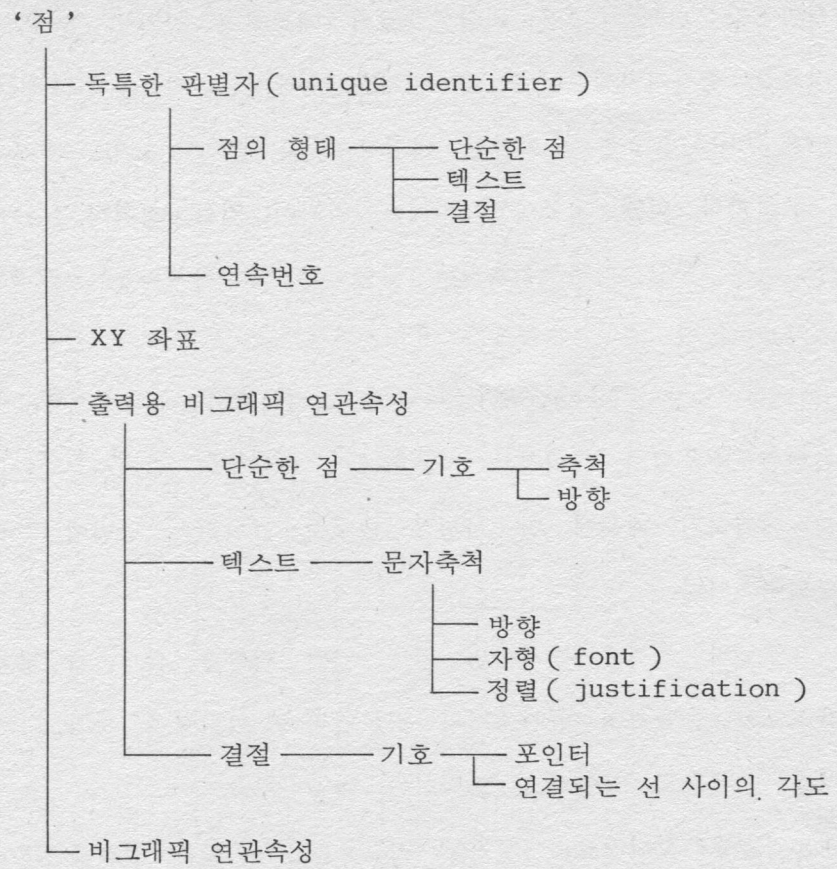
4지수형의 가장 큰 문제점은 4지수 표현이 적용마다 다른 결과가 나타날 수 있다는 것이다.(translation-variant). 즉, 같은 형태와 크기를 가진 두개의 지역이 다른 4지수형으로 될 수가 있다. 따라서 이 방

법은 형태의 분석이나 패턴의 인식에는 적합하지 않다. 그러나, 4지수형은 이질적인 지역을 쉽게 분리해 내고 도형내의 구멍(hole)을 표현할 수 있다. 이 방법은 적용범위나 개발여지가 많은 기술이며, 그 이용에 대한 관심이 점차 고조되고 있다.

2.4.4 벡터(vector)자료구조

벡터식 표현의 목적은 사상을 가능한 한 정확하게 표현하는 데에 있다. 좌표공간을, 래스터 공간과 같이 분할된 것이 아니라 위치, 길이, 차원을 정확하게 표현할 수 있는 연속적인 것으로 가정한다. 그러나 컴퓨터 용량에 제한이 있고, 또한 벡터 표현기기가 래스터 기기보다 해상력이 훨씬 좋지만 기본적 스텝(basic step size)을 내재하고 있기 때문에 엄밀한 의미에서는 연속적이 아니다. 벡터 시스템에서 지리사상은 점•선•면의 조합으로 표현된다.

점사상은 하나의 XY좌표로 위치가 정의되는 지리적 사상이라고 할 수 있다. XY좌표 이외에도 이 '점'이 무엇이며 어떤 속성을 가지고 있는가를 나타내는 정보가 추가되어야 한다. 예를 들어 '점'은 다른 정보와 관련없이 기호화 될 수 있다. 자료의 레코드는 기호와 그 크기, 기호의 방향에 관한 정보를 포함해야 한다. '점'이 문자일 때에는 자료 레코드는 문자의 철자, 활자의 형태(font), 정렬(좌측정렬, 우측정렬, 중앙정렬), 축척, 방향 등에 관한 정보뿐 아니라 다른 비그래픽 속성과의 관련 정보도 포함해야 한다. 〔그림11〕은 '점'사상 자료구조의 예를 보이고 있다.



〔그림 11〕 점 사상의 벡터자료구조

선 사상은 2개 이상의 좌표로 구성되어 있는 선형 (linear feature) 이다. 가장 단순한 선은 시작점, 끝점 (2개의 XY좌표), 그리고 기호를 지시하는 레코드로 구성되어 있다. 기호를 지시하는 파라미터 (display symbol parameter)는 그래픽 출력장치 (disply device)에 선분이나 점선을 불러낸다. '아크 (arc)', '체인 (chain)', '스트링 (string)' 은 연속적인 복잡한 선을 묘사하는 다수 (n개)의 XY좌표 집합이다. 두 개의 XY좌표로 정의되는 선분이 짧을 수록, 또는 XY좌표의 수가 많을 수록 체인은 복잡한 곡선에 가까워진다. 한편, 처리시간이 더 소요되지만

특정한 한개의 값(number)을 저장함으로써 자료저장 공간을 절약할 수 있다. 선형자료를 출력기기에 보낼 때 이값을 통해서 기기 운용 루틴(display drive routine)은 수학적인 내삽함수(B spline)가 좌표계에 적용되도록 지정한다. 이때 좌표계는 이미 기억되어 있는 것이다. '점' 또는 단순한 선과 같이 체인도 출력에 관한 정보를 지시하는 레코드와 함께 저장될 수 있다.

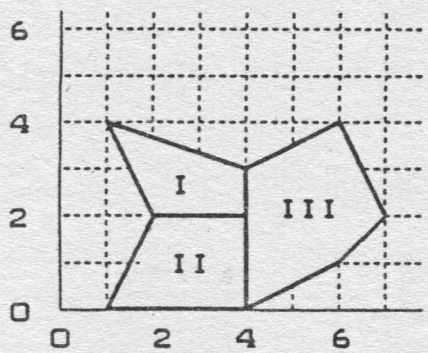
다각형(polygon 또는 region)은 벡터 데이타베이스에서 여러가지 방법으로 표현될 수 있다. 지리정보 시스템에서 대부분의 주제도는 다각형을 사용하여 제작하기 때문에 면 사상을 다루고 표현하는 방법에 많은 관심이 집중되어 있다.

다각형 자료구조의 목적은 면사상의 형태, 인접 다각형, 위계, 즉 위상적 속성을 효율적으로 운용하여 공간조직의 기본단위인 다각형이 가지고 있는 특정한 속성을 주제도로 표현하고 조작하는 데 있다.

다각형 네트워크를 지리자료화 시키는 데는 다음과 같은 전제조건이 있다. 첫째, 지도에서 단위 다각형(region)은 각각 독특한 모양, 둘레, 그리고 면적을 가진다. 여기에는 래스터 방식에서와 같이 기본적인 표준 단위(화소의 크기)가 없다. 규칙적으로 배열된 미국의 시가계획도에서 시가구역의 전부 혹은 일부가 정확히 같은 모양과 크기를 가졌다고 볼 수 없다. 토양도나 지질도에서도 지도단위(map unit)의 형태나 크기가 같은 경우는 거의 없다. 둘째, 지리적 분석을 위해서 자료구조가 하계망(stream network)에서와 같이 각각의 다각형에 인접한 다각형을 기록할 수 있어야 한다. 세째, 주제도에서 다각형은 모두 같은 수준(level)이 아니다. 호수에 있는 섬을 예로 들면 호수라는 큰 다각형안에 섬이라는 작은 다각형이 있는 경우에 두 다각형의 수준은 다른 것으로 처리된다.

〔다각형 구조〕

다각형 구조에서 각 데이타 레이어는 〔그림 12 〕와 같이 다수의 다각형으로 구성되어 있다. 다각형은 다각형을 둘러싼 선분의 좌표로 부호화 되어있고, 다각형은 각각 저장되어 있다. 이 구조에서는 참조 시스템(referencing system)이나 위상관계(topological organization)가 명시적으로 정의되어 있지 않다. 토지이용 소유권과 같이 속성자료는 좌표자료와 함께 수록될 수 있다. 이 구조에서는 두개의 인접한 다각형이 공유하고 있는 변이 두번 정의되는 등 저장공간의 이용이 비효율적이다.



Polygon I:	Polygon II:	Polygon III:
1,4	2,2	6,4
4,3	4,2	7,2
4,2	4,0	6,1
2,2	1,0	4,0
		4,2
		4,3

〔그림 12 〕 다각형 구조

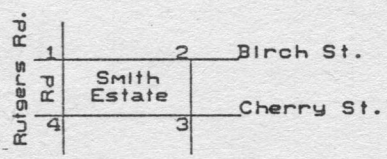
〔DIME 구조〕

명시적 위상 관계를 지리 자료구조에 처음 도입시킨 것은 미국 통계국의 DIME(Dual Independent Map Encoding) 시스템이다. DIME 데이타 화일의 기본 요소는 두개의 끝점으로 정의된 선분(line segment)이다. 복잡한 곡선은 이 선분을 연결시켜 표현한다. 각각의 선분은 양쪽 끝부분에 각

각 한개의 포인터를 가지고 결절과 연결되어 있으며, 또한 다각형을 조립하기 위한 부호(code)를 가지고 있다. 〔그림 13 〕은 간단한 DIME 구조를 나타내고 있다.

각 선분에는 선분 이름, 선분이 시작되고 끝나는 점(from node/to node), 그리고 선분의 좌우에 위치한 다각형이 저장된다. 헤더를 이용하면 속성정보가 쉽게 첨가할 수 있다. 그림에서는 헤더, 우편번호(zip code), 선분이름(가로명)을 이용하여 지번(address)을 확인할 수 있다.

DIME에는 결절에서 만나는 선분을, 혹은 하나의 선분에서 인접하는 선분을 역으로 검색(point-back)할 수 없기 때문에 복잡한 굴곡을 갖는 선분을 처리하거나 다각형의 경계선을 맞추는 것은 어렵다. 같은 가로명을 갖는 선분이 횡단선으로 단절될 때 많은 컴퓨터 시간이 소요된다. 미 통계국에서는 DIME 화일의 오차를 찾기 위해서 ARITHMICON이란 편집 프로그램을 사용하고 있다. 또한 이 자료구조에서는 복잡한 곡선을 다수의 직선으로 구성하고 있기 때문에 데이타의 양이 크다.



Header Items: Groups of Segments

Header Number	Zip Code	Area Code	Tract
1000	93106	805	14
1001	93117	805	14

Segment Codes: Each Line Segment

Segment Name	Nodes From	Nodes To	Polygon Left	Polygon Right	Addresses Left Low	Addresses Left High	Addresses Right Low	Addresses Right High	Header
Birch St	1	2	-	Smith Est	101	175	102	178	1000
Cherry St	3	4	-	Smith Est	103	177	104	180	1000
Rutgers Rd	4	1	-	Smith Est	8602	8686	8603	8685	1000

Node Locations:

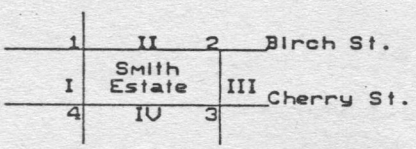
Node	Easting	Northing
1	127,251	1,340,600
2	127,352	1,040,601
3	127,350	1,040,584
4	127,256	1,040,502

〔그림 13〕 DIME 화일구조

〔아크-노드(arc-node) 구조〕

아크-노드 구조에서는 자료를 점-선-다각형의 위계조직으로 구축한다. 〔그림 14〕에서는 지도내의 노드를 x, y 좌표로 정의하고, 정의된 노드를 사용하여 아크를 정의한다. 다각형은 다시 그를 둘러싸고 있는 아크로 정의된다. 이 자료구조에서는 점·선분·다각형이 한번씩 저장되기 때문에 매우 비효율적이다. 다양한 속성정보를 쉽게 저장할 수 있고,속성정보들이 도형과 명시적으로 연결될 수 있다는 장점을 갖고 있다. 즉, 교통통제 장치에 관한 속성정보가 노드에 저장되는 등 적절하게 처리되고 있다.

한편, 아크의 속성가운데 길이와 다각형의 속성가운데 둘레, 면적은 속성자료에 기록되어 있지 않아도 좌표를 이용하여 계산해 낼 수 있는 것들이다. 이들 자료를 자주 사용하는 경우에는 사용할 때마다 다시 계산한다는 것은 비경제적일 수 있으나, 자주 사용하지 않는 경우에는 용량의 측면에서 비경제적이다. 따라서, 프로젝트의 특성에 따라 설계자가 그 field의 설정을 결정해야 한다.



Nodes:

Number	Easting	Northing	Traffic Control	Crosswalk
1	126.5	578.2	Light	Yes
2	218.6	581.9	Sign	Yes
3	224.2	470.4	None	No
4	129.1	471.9	Sign	No

Arcs:

Number	From	To	Length	Pavement Condition	Lanes
I	4	1	106.3	Good	4
II	1	2	92.2	Fair	4
III	2	3	111.6	Fair	2
IV	3	4	95.1	Fair	2

Polygons:

Name	Owner	Arcs	Perimeter	Area	Zoning
A34	J. Smith	I,II,III,IV	405.2	10203	R-4
A35	R. White	III,VI,VII,XI	478.1	11562	R-4

〔그림 14〕 아크-노드 자료구조

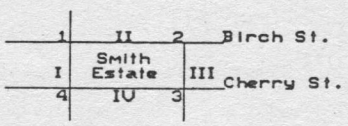
〔관계적 자료구조〕(relational structure)

관계적 자료구조는 아크-노드 구조와 매우 유사하다. 아크-노드 구조에서는 속성자료가 위상자료와 함께 저장되는 데 반해, 이 구조에서는 속성정보는 따로 관리된다.

〔그림 15〕에서 살펴보면 위상자료는 아크-노드 구조의 것과 동일하게 조직되어 있는 것을 알 수 있다. 속성정보는 테이블에 정리되어 있는데, 이 테이블을 릴레이션이라고 한다. 릴레이션의 행(rows)은 레코드(record) 또는 튜플(tuple)이라고 하고 열(column)은 필드 또는 속성이다.

위상자료와 테이블을 연결시키는 것은 두 자료가 공유하고 있는 필드(속성)로 그림에서는 Number와 Name이다. 이들 필드를 키(key)라고 한다.

속성자료를 독립된 화일로 관리하기 때문에 화일의 크기를 늘이거나 화일수를 증가시키면 속성자료를 쉽게 첨가시키고 관리할 수 있다. 특히, 일반목적의 데이타베이스 관리시스템이 관계적 자료구조를 지원하기 때문에 GIS에 연결시켜 사용하면 비용을 절감하고 융통성을 확장시킬 수 있다. dBASE Ⅳ, Oracle 등이 GIS에서 많이 이용되고 있다.



Nodes:

Number	Easting	Northing
1	126.5	578.2
2	218.6	581.9
3	224.2	470.4
4	129.1	471.9

Number	Traffic Control	Crosswalk
1	Light	Yes
2	Sign	Yes
3	None	No
4	Sign	No

Arcs:

Number	From	To
I	4	1
II	1	2
III	2	3
IV	3	4

Number	Length	Pavement Condition	Lanes
I	106.3	Good	4
II	92.2	Fair	4
III	111.6	Fair	2
IV	95.1	Fair	2

Polygons:

Name	Arcs
A34	I,II,III,IV
A35	III,VI,VII,XI.

Name	Owner	Perimeter	Area	Zoning
A34	J. Smith	405.2	10203	R-4
A35	R.White	478.1	11562	R-4

〔그림 15〕 관계적 구조

〔수치선분 그래프(DLG) 구조〕

미 내무성의 USGS는 수치지도자료 확보를 위해 많은 예산을 투자해 왔는데, 특히 이러한 자료의 국가표준을 개발해 오고 있다. 이가운데 수치선분 그래프(Digital Line Graph, DLG)로 알려진 표준 벡터 화일 포오맷이 있다. DLG의 자료원으로는 USGS 7.5 min(1:24,000) 및 15 min (1:62,500) 지형도가 사용되었다. 또한 1:2,000,000 축척의 지도계열을 기초로 DLG 화일이 제작되었다.

DLG화일의 자료내용은 서로 다른 주제 레이어로 구성되어 있다. 레이어1은 주경계 및 행정경계 등을 포함한 경계선, 레이어2는 하천·호수등 수자원, 레이어3은 교통망으로 구성되어 있다. 레이어4는 공공토지 조사 시스템(Public Land Survey System, PLSS)을 기초로 구성되어 있다.

DLG의 기본적 자료요소는 DIME, Arc-Node, 관계적 구조와 유사하다. 결절(node)은 선분의 종점, 또는 선분의 교차점을 나타내며, 선분의 특징 또는 형태를 표시하기 위해서 점(point)이 부가되기도 한다. 선분은 시점과 종점이 있고, 방향과 선분의 좌우에 있는 다각형(지역)을 지정할 수 있다.

또한 길이가 0으로 정의되는 특수한 degenerate line을 사용하여 내용이나 특징을 점으로 정의할 수 있다.

다각형은 선분으로 둘러싸이며 한 지점을 통하여 다각형의 특징을 나타낸다.

점·선·다각형이 위상관계와 위치를 나타내는데 사용되며, 각 요소의 속성정보를 부호화시키는 데에는 코드 시스템을 사용한다. 속성코드는 메이저 코드(major code)와 마이너 코드(minor code)로 조직되어 있다. 메이저

코드는 세자리 숫자로 구성되어 있는데 처음 두자리는 일반적 범주를 나타내고 세번째 자리는 부가적 특징을 나타낸다.

메이저 코드	범 주
020	기복 또는 고도
050	수 문
070	지표피복상태

마이너 코드는 네자리로 구성되는데, 첫자리는 0이다. 다음 세자리는 다음과 같다.

마이너 코드	범 주
001-099	결절(nodes)
101-199	다각형(areas)
201-299	선분(lines)
301-399	degenerate lines
401-499	일반 목적 코드(general-purpose)
601-699	설명 코드(descriptive code)

일반목적(general-purpose) 코드는 사상의 크기와 위치에 따라 점,선 또는 다각형으로 수치화될 수 있는 사상에 할당한 것이고 설명코드는 부가적 정보를 보충하기 위한 것이다.

DLG 자료화일에는 헤더 레코드가 처음 기록되는데, 헤더 레코드에는 화일 작성 일자, 투영법, 좌표 시스템, 화일내에 수록된 점·선·다각형의 갯수 등의 정보가 수록되어 있다. 결절의 자료레코드에는 위치등을 나타내는 설명

정보, 메이저 및 마이너 속성 코드, 문자(text string)이 포함되어 있고, 선분의 자료레코드에는 시점과 종점 좌우의 다각형 등과 같은 설명정보 (description), 선분을 나타내는 순차적인 x, y 좌표, 속성 코드, 그리고 문자가 수록되어 있다.

또한, 자료의 정확도에 관한 정보를 기록한 레코드가 비치되어 있는 경우도 있다. 벡터 자료구조의 개발과정을 고려할 때, DIME과 DLG는 같은 계열이며, 이들을 바탕으로 1990년 미국 센서스에 적용시킨 TIGER(Topologically Integrated Geographic Encoding and Referencing) 시스템이 개발되어 왔다.

2.4.5 벡터와 래스터 방법의 비교

래스터와 벡터방법은 지리정보의 모형화에서 뚜렷하게 구별되는 접근방법이다. 1970년대까지만 하더라도 두 방법은 서로·조화를 이룰 수 없는 대안이었으나, 최근 하드웨어와 소프트웨어의 발달로 상호변환이 쉬워졌고 하나의 시스템에서 두 방법을 동시에 수용할 수 있게 되었다. 이 두 방법의 장·단점을 비교하면 〈표1〉과 같다.

〈표1〉 **래스터와 벡터방법의 비교**

(벡터)

장 점

a) 현상적 자료구조를 잘 표현할 수 있고,

b) 자료구조가 축약되어 있으며,

c) 네트워크 연계로 위상관계를 잘 구축할 수 있고,

d) 그래픽의 정확도가 높고,

e) 위치와 속성의 검색, 갱신 그리고 일반화가 가능하다.

단 점

a) 자료구조가 복잡하고,

b) 지도중첩이 복잡하고,

c) 각 단위가 다른 위상형태를 갖기 때문에 모의(simulation)가 어렵고,

d) 표시기기(VDU)나 도화기가 비싸며,

e) 다각형 내의 공간분석이나 필터링은 불가능하다.

(래스터)

장 점

a) 자료구조가 간단하고,

b) 지도중첩이나 원격탐사자료와 연결시키기가 쉬우며,

c) 다양한 공간분석을 쉽게 할 수 있고,

d) 공간단위가 같은 크기와 형태를 갖기 때문에 모의가 쉽고,

e) 기술자체가 비싸지 않으며 발달속도가 빠르다.

단 점

a) 그래픽 자료의 양이 많으며,

b) 자료를 줄이기 위해 셀을 크게 하면 현상적으로 파악할 수 있는 구조나 정보를 잃어버릴 가능성이 크며,

c) 출력의 질이 나쁘며,

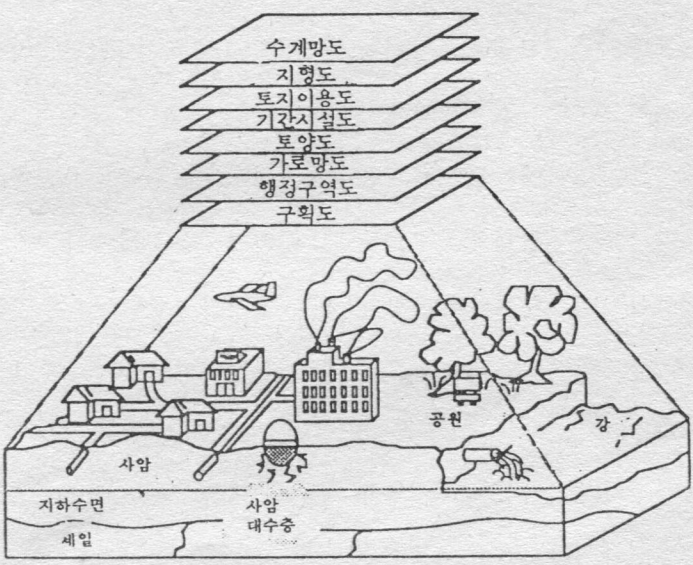
d) 네트워크 연계를 구축하기 어렵고,

e) 특수한 알고리즘이나 하드웨어를 사용하지 않으면 투영변환에 많은 시간이 소요된다.

2.5 GIS의 기능

GIS는 컴퓨터 기술을 이용하여 기타자료의 입력, 보관, 처리 및 출력을 보조하는 정보관리체계이다. 이러한 간단한 정의는 GIS를 단지 하나의 도구로 본 것인데, 만일 GIS를 하나의 시스템이 갖추어야 할 구조와 조직을 완비한 것으로 본다면 GIS는 컴퓨터의 하드웨어와 그것을 지원하는 소프트웨어를 운용하여 상호관련되는 기관과 전문인력을 완벽한 형태로 결합시키는 지리자료 처리 시스템이며, 또한 정보세계를 지원하는 계획을 세우는 일환이라고 볼 수 있다. GIS는 각종의 지리자료를 분산배치하고 표준적인 좌표체계를 통하여 하나의 완벽한 배열로 정리 및 결합시키기 때문에, 대상지역의 각종 자연 및 인문적 특성을 지리적 데이타베이스로 구축해 낼 수 있다〔그림 16〕.

일단 데이타 베이스가 구축되면, 사용자는 프로그램을 이용하여 분석과 시뮬레이션(Simulation)을 수행하게 된다. 이러한 작업의 결과로 얻은 합리적인 해답은 의사결정 과정의 보조자료로 활용한다.



〔그림 16〕 GIS와 레이어 구조

2.5.1 지리자료의 수치화

자료의 획득은 필요한 자료를 조사·확인하고 수집하는 과정으로 다양한 방법이 동원된다. 프로젝트에 필요한 자료와 자료원을 조사하고 조사방법도 정리한다.

야외조사 또는 방문·우편·설문조사 등을 통한 1차 자료가 이상적이지만, 비용과 시간이 많이들고, 때로는 불가능한 경우도 있기 때문에 2차자료의 사용이 불가피할 때가 많다. 지도·항공사진·인공위성 이메이지·타기관의 측량자료·간행자료 등이 많이 사용된다.

자료의 획득과정에서는 소요되는 시간과 경비의 계획이 철저해야 하며, 자료의 정확도·정밀도 등에 세심한 배려가 뒤따라야 한다.

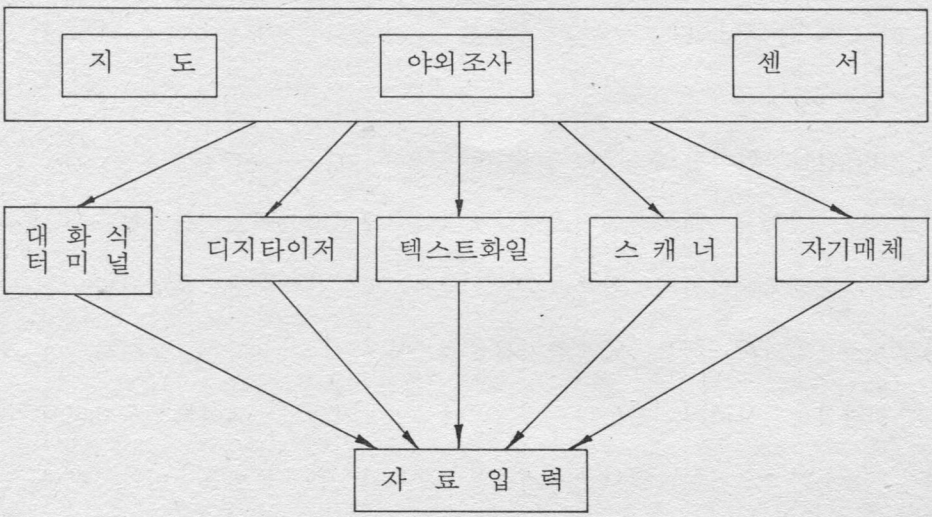
전처리는 GIS에 입력될 수 있는 형태로 자료를 처리하는 과정을 말하는데, 구체적으로는 데이터 포맷의 변환과 데이타베이스내의 위치지정을 가리킨다.

지도·항공사진·조사보고서 등으로 부터 정보를 추출하여 컴퓨터가 읽을 수 있도록 부호와(Encoding)하여 컴퓨터 데이타베이스에 기록하는 것 등이 데이타 포맷의 변환에 포함된다. 이 과정은 시간과 비용이 많이 들고 오류가 끼어들 수 있는 기회가 많다.

한편, 데이타베이스내에 오브젝트를 기록하고 위치를 지정할 때에는 일관된 체계에 따르도록 한다. 예를들면 레이어 시스템에서는 자료가 속한 레이어에 따라서 그 특성을 파악할 수 있기 때문이다. 또한 이 과정에 자료의 질을 통제할 수 있는 기준을 포함시키면 데이타베이스의 효용을 높일 수 있다.

지리자료의 수치화 작업은 기존의 처리 또는 도면자료와 그 속성자료를 변환시켜 컴퓨터가 받아들일 수 있는 형식으로 변환시키는 것을 가리킨다.

도형을 수치화시키는 일반적인 도구로는 디지타이저(Digitizer), 스캐너(Scanner), 비디오 카메라(Video Camera)등이 있으며, 지리공간이 가지는 고유한 속성은 키보드를 통해 입력시키는 경우가 대부분이지만 스캐너로도 속성자료를 입력시킬 수 있다〔그림 17〕.



〔그림 17〕 자료의 입력

수치자료화의 속도에 있어서는 비디오 카메라가 가장 빠르며, 그 다음이 스캐너와 디지타이저의 순이다. 반면에 가장 느린 것은 키-보드를 이용하는 경우이다. 자료의 질에 있어서는 일반적으로 스캐너가 가장 좋으며 그 다음이 디지타이저, 키-보드의 순이고 비디오 카메라가 가장 떨어지나, 자료의 종류나 이용목적, 그리고 경비와 소요시간 등을 고려하여 입력수단을 결정해야 한다.

이들 이외에도 원격탐사로 얻는 화상자료가 있는데, 이것은 지면의 상황을 이미 수치화시킨 것이기 때문에 직접 컴퓨터에 입력시킬 수 있다.

수치화하여 입력시킬 수 있는 주요한 지리정보로는 다음과 같은 것들이 있다.

a. 원격 탐사자료

b. 토지이용자료

c. 지적자료

d. 인구 및 주택자료

e. 지형자료(DEM, DTM)

f. 지질자료

g. 전신·전화자료

자료관리 기능은 데이타베이스의 제작과 데이타베이스에의 접근을 통제하며, 이 기능은 데이타의 입력, 갱신, 삭제 및 검색으로 수행된다. 최근 데이타베이스 관리시스템은 급격한 진보를 이루어 사용자는 오히려 이 기능에 큰 관심을 보이지 않는다. 그러나 이 기능이 제대로 수행되지 못할 때에는 사소한 오퍼레이터나 기계상의 오류가 큰 문제를 야기시킨다. 또한 이 기능에서 안전관리(Security management)도 이루어져야 하는데, 사용자의 수준에 따라 시스템과 데이타베이스의 접근할 수 있는 정도를 통제하는 프로시듀어가 장착되어야만 한다.

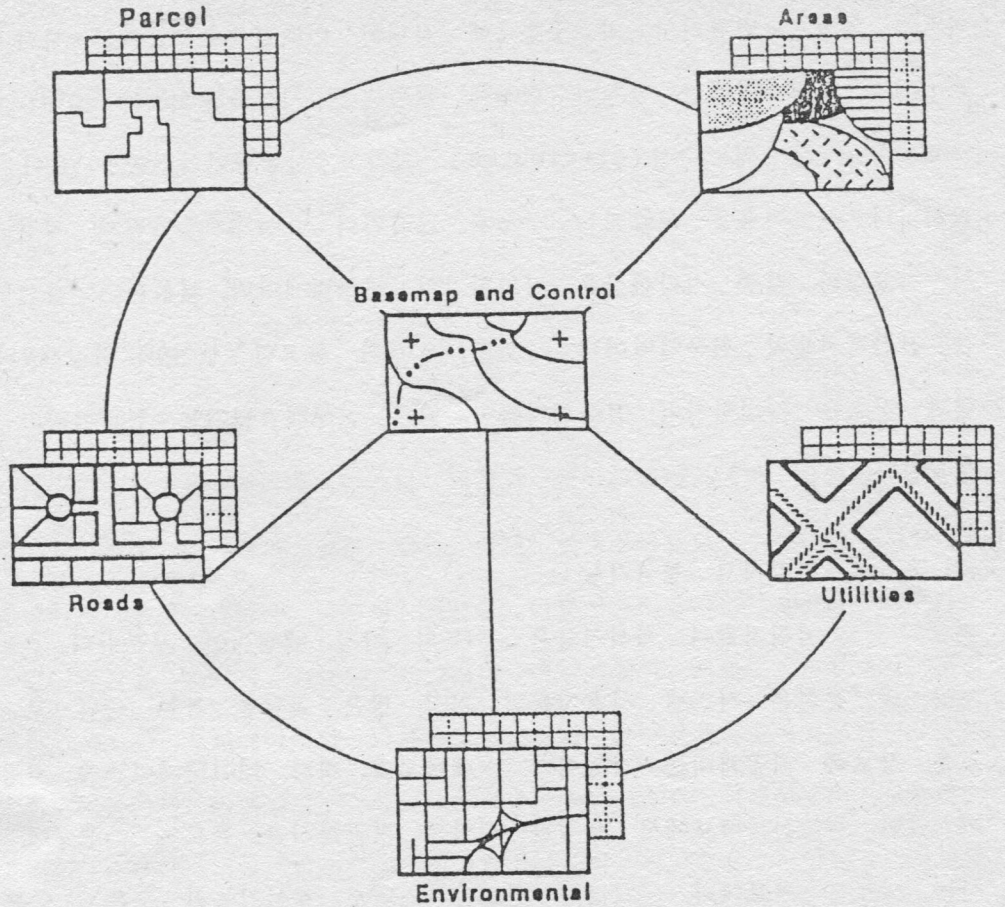
자료관리 시스템(DBMS)은 GIS에서 가장 중요한 부분으로 자료의 저장과 처리 및 분석의 전 과정의 효율성에 직접적인 영향을 미친다. 지리데이타베이스 시스템이란 지리정보의 기록체계를 말하는데, 이것은 지리정보를 효과적인 방식으로 저장하여 사용자에게 적시에 적절한 정보를 제공하기 위한 것이다.

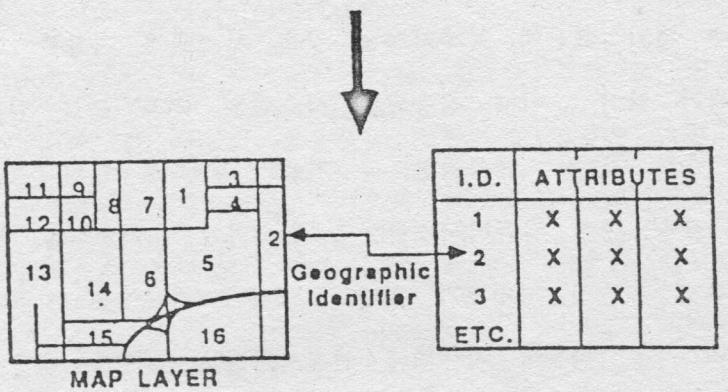
하나의 완벽한 지리적 데이타베이스 시스템이 되기 위해서는, 우선 지표상의 각종 지리현상의 특성이 개념적인 지리자료 모형(conceptual model of geographical data)속에 반드시 포함되고 이러한 개념적인 모형이 지리정보 시스템의 구조속에 포함되도록 구체화되어야 하며, 또한 지리현상

과 그 특성을 컴퓨터의 자료구조속에 적당히 안배할 수 있어야 한다.〔그림 18〕. 개념모형이란 특정한 현상에 관련되는 사상유형(geographic entity), 사상의 속성(attributes) 그리고 사상들간의 관계 등이 결합하거나 상호작용을 일으키는 구조를 가리킨다. 이러한 일련의 과정중에서 지표상의 각종 지리현상은 지리학자의 연구과제이며, 컴퓨터의 자료구조는 전산학자들의 전문영역이다. 또한, GIS의 중요연구분야인 지리자료의모형과 구조는 지리학자와 전산학자들이 함께 연구해 나가는 영역이라고 볼 수 있다.

2.5.2 지리자료 분석기능

지리자료의 분석기능은 GIS의 가장 핵심적인 부분이다. 또한 처리 및 분석은 시스템 사용자들이 가장 많은 관심을 쏟는 측면으로 새로운 정보를 가공해내는 기능이다. 예를들면, 지리 데이타베이스에 저장되어 있던 등고선 정보에서 특정한 지역의 평균경사를 유도하는 경우에 경사는 새로운 정보이다. 그러나 어느 시스템도 분석기능의 모든 종류를 구비할 수 없기 때문에, 특수기능을 갖춘 시스템에 자료를 전송해서 특수분석할 때가 있다. 이때 전송되는 시스템은 GIS 또는 일반적인 통계패키지일 수도 있다. 예를들면, 한 기관의 GIS로 부터 자료를 수치 분석기능을 갖춘 대형 컴퓨터시스템으로 전송하여 수치분석을 하고 다시 원래의 GIS로 전송하는 경우가 많다. 그러나 최근에는 이러한 특수기능을 모듈(module)로 만들어 시스템에 연결시키려는 전략을 채택하고 있는데, 이 전략으로 기존의 시스템을 기초로 분석기능을 점진적으로 확장시킬 수 있다. 일반적으로 geoprocessing이란 GIS의 처리 및 분석기능을 지칭한다. 분석기능을 다음과 같이 분류할 수 있다.





MAP LAYER AND TABULAR ATTRIBUTES LINKED BY GEOGRAPHIC IDENTIFIER

〔그림 18〕 지리자료의 개념모델

A. 공간자료처리 Spatial manipulation

a. 벡터방식과 래스터방식간의 변환 Vector/raster conversion

b. 좌표변환 Transformation

c. 좌표화 Coordination

d. 인근분석 Proximal or neighborhood analysis

e. 면적과 둘레의 계산 Area/perimeter calculation

B. 공간자료 분석 Spatial analysis

a. 다각형의 중첩 Polygon overlay

b. 다각형의 삭제 Polygon dissolve

c. 다각형 내부의 점 Point in polygon

d. 완충구역 및 회랑지역 설정 Buffering and Coridoring

e. 윈도우윙 Windowing

C. 수치지형 분석 Digital terrain analysis

a. 등고선 작성 Contouring

b. 경사와 주향계산 Slope/aspect

c. 하천유역 분석 Watershed analysis

d. 체적 계산 Volume calculation

e. 단면도 작성 Cross Section

f. 3차원 입체 화상 제작 3-D viewing

D. 네트워어크 분석 Network analysis

a. 적정노선 선택 Optimal path selection:route

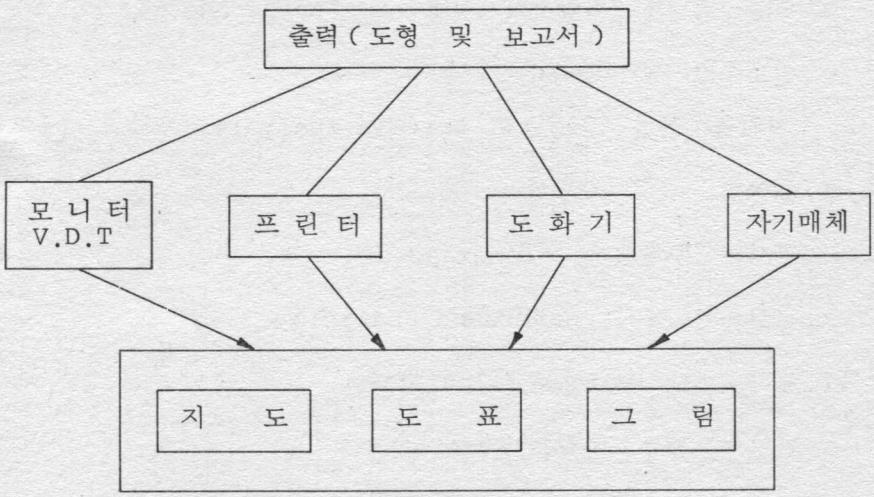
b. 유통량 시뮬레이션 Flow simulation:allocate

c. 시간 및 거리 구분 Time/distance districting:district

2.5.3 출력

출력에는 지역별 인구밀도·인구수와 같은 통계 테이블, 각종 주제도, 그래프 등이 있다. 또한 종이나 마이크로필름, microfiche와 같이 인식할 수 있는 형태와 자기테이프, 컴팩트 디스크등과 같이 컴퓨터가 읽을 수 있는 비가시적인 것이 있다. 또한 모니터에 일시적으로 나타나는 소프트카피(softcopy)와 종이나 필름에 기록한 하드카피(hardcopy)로 나누기도 한다〔그림 19〕.

최근에는 자료의 저장·교환에 국가적 또는 세계적 표준이 점차 적용되면서 자기매체로 출력하는 경향이 늘어가고 있다.



〔그림 19〕 자료의 출력

분석·처리된 지리정보를 사용자에게 정확히 전달하기 위해서는 표현수단(예, 지도)을 잘 설계해야 한다. 지도를 설계할 때는 반드시 지리정보가 표현하는 주제와 대상을 고려하여 정보의 양을 적당하게 조절하고, 사

용자가 적절한 정보를 편리하면서도 정확하게 판독할 수 있게 해야한다. 또한, 지리정보의 표현수단을 설계할 때는 출력시키는 도구의 성능을 고려하여 설계한 의도가 충분히 반영되도록 해야한다. 현재 일반적으로 사용되어지는 도구로는 프린터, 플로터, EPP (정전식 플롯 프린터), 모니터 등이 있다.

색상을 표현하는 데에는 모니터가 가장 아름답고, 그 다음이 Color EPP와 프린터이며, 플로터가 가장 뒤떨어진다. 한편 플롯(Plot)의 질에 있어서는 플로터와 모니터가 가장 뛰어나며, 그 다음이 EPP이며 프린터가 가장 뒤떨어진다. 선형의 벡터자료를 표현하기 위해서는 플로터나 모니터가 적절하며, 반면에 레스터식 자료를 표현하는 데는 프린터나 EPP를 사용하는 것이 좋다. 한편, EPP는 선형 벡터자료의 표현에도 이용할 수 있다.

정보표현(display)을 위한 설계에서 여러가지 도구들의 특성을 혼합시키면 양질의 출력을 사용자에게 제공할 수 있으며, 출력으로 나온 지리정보의 분석결과는 현상해설이라는 응용적인 측면의 기능도 발휘할 수 있다.

현재, GIS는 분산식 처리와 여러 네트워크를 결합시키는 방향으로 발전하고 있다. 이것은 각 부문별로 공정을 나누어 협업하는 방식과 마찬가지로, 각각의 네트워어크는 수치화된 지리자료를 개별적으로 관장하지만, 네트워어크간의 전송기능을 통해 상호 유통되어 결과적으로는 지리자료 자체가 공유되어진다. 컴퓨터 하드웨어나 GIS소프트웨어 역시 이러한 방향으로 발전해가고 있다. 이러한 분산식 처리방식의 활용을 통하여 수년 후에 GIS기술은 각 지방정부의 모든 업무에 보급 이용되어질 수 있을 것이다.

2.5.4 GIS의 개발

GIS의 개발은 일반적으로 타당성의 평가(feasibility assessment), 시스템의 설계 및 구입, 수행, 그리고 운용 및 관리의 4단계로 나눌 수 있다. 각 단계를 추진하는 데에는 많은 노력과 여러가지 활동이 요구된다. 각 단계의 활동은 다음과 같이 분류할 수 있다〔그림20〕.

1. 개념적 시스템 설계(Conceptual System design)
2. 시스템의 구입 및 설치(System acquistion and installation)
3. 데이타베이스의 개발(Database development)
4. 실무에 적용(Application development)
5. 훈련 및 업무수행(Training and service)

GIS개발의 5단계를 간략하게 살펴보기로 하자.

1단계 : 개념적 시스템 설계(Conceptual System Design)

GIS개발의 제1단계로서 GIS개발의 타당성과 시스템의 개념을 설계하고, 실행의 전반적인 계획을 수립하는 단계이다. 이 단계에서 GIS의 경영진, 전문가, 기술진에게 시스템의 목표를 설명하고(orientation), 필요사항을 수집하고 시스템 개발계획을 수립한다. 이 단계에는 다음의 6가지가 주요한 내용이다.

1. 연구계획 오리엔테이션(Project Orientation)
2. 필요사항 평가/타당성 조사(User Need Assessment/Feasibility Study)
3. 데이타, 시스템과 조직의 구비요건(Data, Systems and Organization Requirments)
4. GIS데이타베이스의 개념(GIS Database Concept)
5. 시스템의 개념(System Concept)

GIS Life Cycle	Developoment Tracks and Activities				
	Conceptual Design	System Acquisition and Installation	Data Base Development	Application Development	Training and On-Going Support
1. Feasibility Assessment	• Orientation • Needs Assessment • Requirements • Cost/Benefit				
2. System Design and Procurement	• Database Design • Application Definition • HW/SW Concept • Organization Concept	• HW/SW Specifications • Tender Documents • HW Vendor Contracting • HW/SW Acquisition	• Physical Database Design • Pilot Study	• Application Specification • Prototype Creation	
3. Implementation	• Implementation Strategy • Funding/Financing Strategy	• HW/SW Installation • System integration • System Programming	• Data Standardization • Data Conversion • Database/library Creation • Product Creation	• Application Creation • Documentation	• User Training • Administrator Training • Programmer Training • Application Training
4. Operations and Maintenance		• System Upgrade	• Update Procedure Definition • Service Bureau to Users	• Application Enhancement	• On-Going Education • Contracted Support • On-Site Support • Funding Program

〔그림 20〕 GIS의 개발

6. 실행 계획 (Implementation Plan)

2단계 : 시스템의 구입 및 설치 (System Acqusition)

제 2 단계는 하드웨어, 소프트웨어, 통신기기 등 시스템을 선택, 구입하고 설치한다. 연구기관에서 필요로 하는 연구계획에 맞는 시스템을 선정하고 설치하는데, 이에는 다음과 같은 실행절차가 있다.

1. 하드웨어의 선택과 구입 (Hardware Specification and Acqusition)
2. 하드웨어와 소프트웨어의 최종 결정 (Finalizing Hardware and Software Agreement)
3. 하드웨어 환경 조성 (Preparing Hardware Environment)
4. 하드웨어와 소프트웨어의 설치 (Installing Hardware and Software)
5. 적합성 평가 (Performing Acceptance Test)

3단계 : 데이타베이스의 개발 (Database Development)

제 3 단계는 지리정보 데이타베이스를 구축하는 단계이다. 데이타베이스를 완성하려면 오랜 시간과 많은 준비가 필요하다. 그러나 제 1 단계의 개념 설계를 통해 이는 대부분 해결되지만, 이 단계에서 구체적인 확인이 있어야 한다. 이 단계에서는 실험연구 이후에 완전 자동화를 실행한다. 이에는 다음의 8가지 실행단계가 있다.

1. 시험 설계 (Pilot Study Design)
2. 운용 데이타베이스의 구체적 설계 (Physical Design of the Operational Database)
3. 데이타베이스 점검 (Quality Assurance Program)
4. 자동화 시험 (Pilot Automation)
5. 자료의 수집 (Data Collection)

6. 자료의 편집과 오류제거(Data Compilation and Rectification)

7. 획득의 자동화(Data Automation)

8. 데이타베이스의 구축(Database Creation)

4단계 : 실무에 적용(Design and Development of User Applications)

이 단계에서는 GIS를 이용한 구체적인 실무적용을 계획한다. 이 활동은 GIS의 초기단계에 설정되어야 하지만, 적용될 때마다 여러번 검토되어야 한다. GIS개발의 초기단계에서는 이 단계의 활동을 제3단계와 동시에 수행한다.

1. 업무의 기능적 정의(Functional Definition of the Applications)

2. 업무의 설계(Application Design)

3. 업무의 구체화(Final Application Specification)

4. 업무의 개발(Application Development)

5단계 : 훈련 및 업무수행(Training and Ongoing Support Services)

이 단계에서는 GIS의 기술과 유지, 적용을 위해 사용자가 그 자원과 기술을 개발하는 단계이다. 이 단계는 제4단계와 병행하여 실행하지만 이것이 자리잡은 후에는 시스템의 지속적인 성장에 원동력을 제공하게 된다. 이에는 다음의 5가지 절차가 있다.

1. 훈련(Training)

2. 하드웨어와 소프트웨어의 보강(Augmentation of Hardware and Software)

3. GIS관리(GIS Management)

4. 지속적인 지원(On-site Support)

5. 재원 확보(Determine Financing Mechanisms)

2.6 GIS의 응용

막대한 분량의 정보를 관리하는 정부기관의 경우 소장자료의 80%-90%가 공간분포의 성질을 갖춘 지리자료이다. 따라서, 정부기관이 GIS의 최대 수요자이다. 그러나, 많은 개인기업이 업무의 기획과 추진과정에서 지리정보를 이용하여 문제를 해결하려는 경우가 적지 않다. 간단한 화물 운송노선의 계획에서부터 복잡한 과정을 거치는 공장부지의 선정이나 상점부지의 선정에 이르기까지 GIS는 뛰어난 문제해결 수단이 될 수 있다. 학술연구 부문에서는 공간정보와 긴밀한 상호관련성을 가지는 분야인 지리학, 지질학, 고고학, 토목공학, 수리학, 지역계획, 도시계획, 지적, 부동산, 그리고 조경건축 등에서 GIS를 이용함으로써 연구성과를 향상시킬 수 있다.

GIS이용의 예는 다음과 같다.

* 국토계획
  - 지역계획
  - 도시계획
* 토지 문제
  - 토지이용 계획
  - 토지자원 관리
  - 해안지역 계획
  - 지가 평가
* 자원분석 및 경영관리
  - 자연자원 분석 경영관리
    - 야생동물
    - 해양동물

· 삼림

— 광물자원 경영관리

· 석유

· 경제성 광물

* 환경관리

— 환경자원의 분석 및 관리

— 환경감시

* 수자원 관리

— 유역 관리

— 하천 유량

— 하천 범람치 계측

— 해일 및 파랑의 관측

* 교통운송도로망 계획

— 도로 노선의 선택

— 교통망 계획

— 교차로 계획

* 국방응용

— 군사기지 및 부지 선정

— 지세 분석

— 전술 모의 훈련

— 행군로 계획

— 직사 및 곡사화기 배치

* 사회안전분야

— 경찰 차량, 소방차, 구호 차량 등의 순시 및 이동노선 분석

- 경찰관서, 소방서 부지 선정 및 관할구역 계획
- 소화전 계획 관리
- 삼림 감시탑부지 선정
- 산불예방
- 자연재해 구조

* 공공시설 관리
  - 공원부지 선정 및 계획
  - 가스 상수도 전화 전기관로 배선
  - 가스 저장 및 가압소 등의 입지 선정
  - 학군계획

* 지 형
  - 지질 상황 모의
  - 토목량(절토 및 성토량) 계산
  - 경사도와 주향분석

* 교통 관리
  - 자동차 컴퓨터 지도
  - 최적 운행 노선 선택

* 상권 및 시장분석
  - 산업지구 분석
  - 산업 수요 및 공급

* 기 타
  - 고고학, 고증
  - 부동산 관리(vector data + videography)
  - 인구 및 취업시장 분석

3. GIS의 발전-회고와 전망

3.1 역 사

지도 이용자들은 오랫동안 대량의 지리정보를 수록하고 판독하는 방법에 있어서 많은 고심을 해왔다. 그런데, 1950년대에 컴퓨터의 발명으로 이러한 문제를 해결할 수 있게 되었으며, 지리정보는 정보화시대의 매우 중요한 분야가 되었다.

3.1.1 발전 배경

지리정보 시스템은 다음과 같은 부문의 발전을 통해 성장해 왔다.

a. 컴퓨터 하드웨어 기술의 진보

b. 수치화 도구의 발전과 지리공간자료의 내실화

c. 컴퓨터제도 분야의 발전과 가격하락

d. 데이타베이스 관리 기술의 발전

e. 수치화 및 공간분석 기술의 발전

3.1.2 성장과정

3.1.2.1 1960년대

컴퓨터 설비가 무척 고가였으며 도화(Plot)시스템의 성능에 한계가 있으면서도 가격은 매우 높아, 이 단계에서의 발전은 모두 공공기관에 의해 이루어졌다. 예를 들자면, 캐나다 정부의 GIS와 60년대 말의 미국의 호구조사 기본도 시스템(GBF/DIME FILE)등이 있다. 한편, 이 시기에는 학술계통에서는 대다수가 레스터식 자료분석을 위주로한 GIS를 발전시켰는데, 함부르크 대학의 GRID, 메릴랜드 대학의 MANS등을 꼽을 수

있다. 그후 캐나다의 GIS와 미국의 DIME은 모두 벡터식 자료를 처리할 수 있게 되었으며, 특히 DIME은 공간적 위상관계(Topology)를 수용하고 있었다.

3.1.2.2 1970년대

북미에서 환경문제는 중요한 관심의 대상이었고, 정부는 GIS와 같이 계획입안에 활용할 수 있는 지원도구를 긴급히 물색하였으며, 이를 통하여 환경정보를 파악하고 신속하고도 적합한 해결책을 구하려는 노력이 있었다. 이 단계에서는 컴퓨터와 그래픽 주변기기 시스템의 성능이 향상되고 한편으로는 가격이 하락하면서, GIS는 미국 연방정부나 큰 기관의 전유물에서 벗어나게 되었고 많은 유관기관에서도 GIS를 발전시킬 수 있게 되었다. 그리고, 지리정보연구회의 학술행사가 경험의 교류와 저변확대에 큰 기여를 하였다. 국제지리학회는 두 차례의 지리정보연구회를 개최하였고, 그후에는 Auto-Carto학술회가 국제적인 컴퓨터 지도학회와 GIS학회를 이끌어 오고 있다. 이 단계에서 GIS는 도약을 보였으나, 대부분 래스터 방식을 위주로 하였기 때문에 데이타베이스 시스템을 결여하고 있는 경우가 많았다.

당시의 응용분야는 크게 두 가지로 나누어지는데, 하나는 천연자원 관리 분야로 그 정밀도는 비교적 낮아도 무난하기 때문에 대부분이 래스터방식을 사용하였고, 다른 하나는 토지나 공공시설의 관리 분야로 비교적 높은 정밀도가 요구되기 때문에 점차 벡터 방식의 사용이 증가하게 되었다.

70년대 말에는 GIS사용자가 증가함에 따라, 전문적인 GIS 소프트웨어 회사가 출범하기 시작했고, 이들을 통해 GIS는 더욱 전문화되어 갔다.

3.1.2.3 1980년대

80년대에는 GIS의 공간자료가 위상(topology)의 측면과

데이타베이스 관리 시스템을 포함하게 되었다. 또한, 워크 스테이션(Work station)의 도입으로 공간자료와 함께 다양한 속성 자료를 동일 시스템 내에서 분석할 수 있게 됨에 따라 경비의 절감은 물론이고 우수한 분석 결과를 얻을 수 있게 되었다. 그 밖의 추세로는 개별적인 공간분석 소프트웨어를 모듈(module)로 결합할 수 있는 조합식 도구 박스(Tool Box)의 소프트웨어 생산이다. 개별적 목표의 데이타베이스 역시 점차로 다용도의 데이타베이스로 통합되고, GIS의 상품화와 다용도화가 급속이 이루어지고 있다.

3.2 발전 전망

이미 개발 소개된 GIS의 다용도화와 상품화 추세 이외에도, GIS의 기본적 기능과 응용의 측면에서 보다 많은 새로운 발전이 있을 것이다.

3.2.1 기본적 기능의 측면

* 입력—측량 및 수치화 도구의 발전은 물론이고, 원격탐사정보는 더욱 다양하고 정확해지며 정보제공 또한 신속해질 것이다.
* 관리—신속한 운용과 계산에 더욱 적합한 공간자료구조의 연구, 인공지능을 이용한 정보상담 방식의 개발, 반자동 정보관리 시스템으로 오차의 해결에 큰 진보를 기대할 수 있을 것이다.
* 분석응용—공간분석 능력의 증강과 아울러 모형화 기능이 강조되어 GIMS(Geographic Information Modelling System)로 이행될 것이다.
* 출력—더욱 다양한 형식을 통해 보다 효과적인 표현방식이 개발되어질 것이며, 자동지도제작 설계분야와 결합될 것이다.

3.2.2 응용측면

* 지리 데이타베이스—더욱 완벽하고, 신속한 운용을 지원해 줄 수 있을 것이다.

* 공간분석—real 3-D의 개발 등으로 응용분야가 더욱 넓어지고, 모형화가 강조될 것이다.

* 인식증가—보다 많은 사용자가 GIS가 제공하는 편리함과 효용성을 활용하고자 할 것이다.

4. 미 조사 통계국의 TIGER시스템

4.1 序

미 조사 통계국은 1981년 주요한 목표를 설정하였는데, 1990년 인구 및 주택 센서스에서 자료조사·통계표 작성(Tabulation)·조사결과 배포에 수반되는 모든 지도적·지리적 처리를 자동화시킨다는 것이었다. 이것은 센서스에 필요한 지도·주소참조화일(address range reference file)·지리사상간의 관계 화일(geographic relationship file)을 전혀 새로운 방법으로 제작하는 것이며, 이를 위해서 미국내의 모든 가로 및 도로, 도로명 및 가로 명칭, 주소, 철도, 강·호수 등과 같은 수자원사상(hydrographic feature), 경계선, 지명, 지역코드(geographic code)등을 전산화하여 하나의 데이타베이스를 통합하는 작업이었다. 그. 결과로 나타난 것이 GIS기술이 성취한 것 가운데 가장 성공적인 TIGER(Topologically Integrated Geographic Encoding and Referencing)화일이다. TIGER시스템 구축에는 약 3억 달러의 경비가 소요되었다.

전통적인 방법으로는 1990년 센서스를 감당할 수 없다는 판단으로 TIGER시스템을 제작하게 되었다. 다음은 그 이유를 정리한 것이다.

1. 사회가 다원화되고 정보사회로 진입하면서 대중의 정보요구는 더욱정교한 계수방법(enumeration methodology)을 요구하게 되었다. 이를 만족시키기 위해서는 조사단위를 더욱 세분하고 조사항목을 증가시켜야 한다.
2. 따라서 더 풍부하고 정확한 경계선 정보(boundary information)가 필요하다.
3. 이 정보는 조사구 요도에 적절하게 나타내야 한다. 30만의 계수관이

조사에 지참할 요도는 약 130만 도엽으로 도엽당 5～6장을 제작한다면 약 700만장의 요도가 짧은 기간에 제한된 장비와 인력을 통하여 마련되어야 한다.

4. 1970년부터 실시된 우편조사법(census-by-mail method)을 활성화시키기 위해서는 지리코드(geographic code)에 주소를 연결시켜야 한다.

4.2 TIGER/database file

데이타베이스의 개념적 설계에는 이용자의 요구 혹은 필요사항을 잘 살피고 그것이 적절하게 지원되도록 구조화시키는 단계가 필수적이다.

TIGER 데이타베이스는 자료수집과 통계작성에 적절하도록 지리적으로 구조화하고, 주소를 지리적으로 부호화(geocoding addresses)시키고, 통계국의 조사구요도 및 발간통계도(product map)의 자료원이 되도록 설계되었다. 이 데이타베이스는 개념적으로는 하나의 화일이지만, 군(county)별로 분할(partition)되어 있는 등 실제로는 TIGER county partition 화일, GEO-CAT (Geographic Catalog of Political/Statistical area)화일, national partition 화일, 그리고 잠정적인 워크 화일의 4가지 종류로 구성되어 있다.

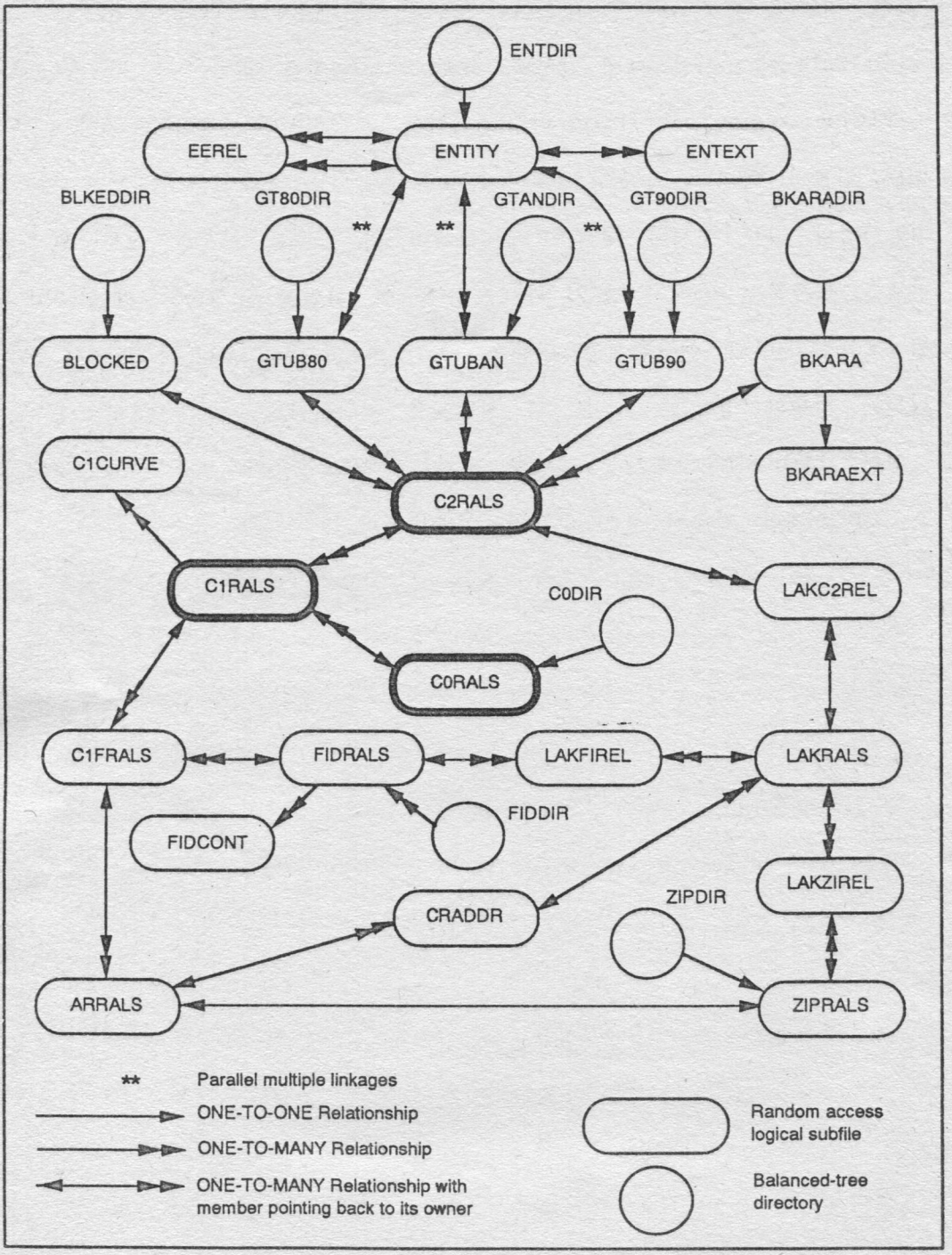
워크화일은 화일을 제작하고 갱신할 때 사용하는 일시적인 것이며, National partition 화일은 모든 군(county) 경계선의 좌표와 partition file identifier를 가지고 있어 참조화일의 역할을 한다.

GEO-CAT 화일에는 지리사상의 현재 및 과거의 정보, 명칭, 지리사상과 county partition file 사이의 관계 등이 수록되어 있고, 각 사상에 지리사상 코드(geographic entity code)를 부여하는 master referen-

ce의 역할과 함께 지리적 명칭이나 지리코드를 가지고 county partition file을 조사할 때에 rapid reference역할을 한다.

TIGER county partition file은 전체 TIGER 데이타베이스의 키(key)이며, 자료의 양이나 화일의 수에서 압도적이다. 이 화일에는 처리된 모든 지리사상의 좌표·코드와 TIGER 데이타베이스의 구조를 이루고 있는 위상관계를 수록하고 있다. 〔그림 21〕과 〈표2〉는 이 화일을 구성하고 있는 부화일(subfile)의 연결관계와 내용을 나타내고 있다. TIGER 시스템에서 출력되는 지도의 정보는 대부분 이 화일에서 지원된다.

한편, TIGER데이타베이스 화일의 자료원(data source)의 주요한 부분은 USGS에서 제공한 1:100,000의 DLG이다.



〔그림 21〕 하나의 **TIGER county patition** 화일내의 부화일

〈표 2〉 TIGER county partition subfiles

Name	Identification
ARRALS	Address Ranges
BKARA	1990 Block/Address Register Area (ARA)
BKARADIR	1990 Block/Address Register Area (ARA) Directory
BKARAEXT	1990 Block/Address Register Area (ARA) Extension
BLKEDDIR	1980 Block/Enumeration District Directory
BLOCKED	1980 Block/Enumeration District
C0DIR	0-Cell Directory
C0RALS	0-Cell (Zero-Dimensional Topological Cell)
C1CURVE	1-Cell Curvature
C1FIRALS	1-Cell/Feature Identifier Relationship
C1RALS	1-Cell (One-Dimensional Topological Cell)
C2RALS	2-Cell (Two-Dimensional Topological Cell)
CRADDR	Cross-Reference Addresses
EEREL	Entity-to-Entity Relationship
ENTDIR	Geographic Entities Directory
ENTEXT	Geographic Entities Extension
ENTITY	Geographic Entities
FIDCONT	Feature Identifier Continuation
FIDDIR	Feature Identifier Directory
FIDRALS	Feature Identifier
GT80DIR	1980 Geographic Tabulation Unit Base Directory
GT90DIR	1990 Geographic Tabulation Unit Base Directory
GTANDIR	Ancillary Geographic Tabulation Unit Base Directory
GTUB80	1980 Geographic Tabulation Unit Base
GTUB90	1990 Geographic Tabulation Unit Base
GTUBAN	1990 Ancillary Geographic Tabulation Unit Base
LAKRALS	Landmarks, Areas, Key Geographic Locations
LAKxxREL	Relationships between the LAKRALs and the C2RALS, ZIPRALS, or the feature FIDRALS
ZIPDIR	ZIP Code Directory
ZIPRALS	ZIP Code

4.3 Cartographic Extract

TIGER데이타베이스는 지리사상을 0-cell(point), 1-cell(line), 그리고 2-cell(area)로 분류하여 저장하고 있기 때문에 이 자료에서 지도를 출력하기 위해서는 지도오브젝트(cartographic object)로 조직해야 한다.

또한 하나의 지리사상을 지도 오브젝트로 정리하기 위해서는 여러단계의 처리를 거쳐야 하고 여러개의 화일을 참조해야 한다. 예를 들면 호수를 지도화하기 위해서는 8개의 화일에서 정보를 가져와야 한다.

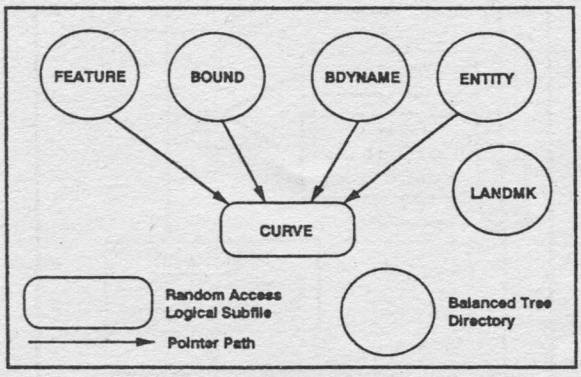
특히, 조사구 요도 제작에서 같은지역을 나타내지만 축척·내용·설계가 약간 다른 여러종류의 지도를 제작할 때 그때마다 이와같이 복잡한 절차를 거쳐야 하는 불편함이 문제가 된다.

이러한 장애를 제거하기 위해 제안된 것이 cartographic extract라고 하는 2차 데이타베이스다. Cartographic extract는 TIGER database에서 정보를 추출하여 구축한 것으로 〔그림22〕, 지도제작프로그램이 좌표스트링을 즉각 불러낼 수 있도록 구조화되어 있다.

Cartographic Extract를 이용하여 조사구 요도 제작은 완전 자동화되어 있고, batch processing으로 수행되고 있다. 지도의 축척, 지명 부착, 삽도(insert map)의 크기 및 축척, 도엽설계(map layout)등의 결정이 자동화되어 있다〔그림23〕.

예를 들면, 축척은 지역내의 지리사상 밀도(feature density)를 정량화하여 결정하는데, 조사구 요도의 축척은 1:5,000에서 1:70,000까지 500을 단위로 증감한다.

한편, 조사구 요도를 계산할 때에는 TIGER database의 수정을 거쳐서 cartographic extract를 갱신하도록 되어 있다.

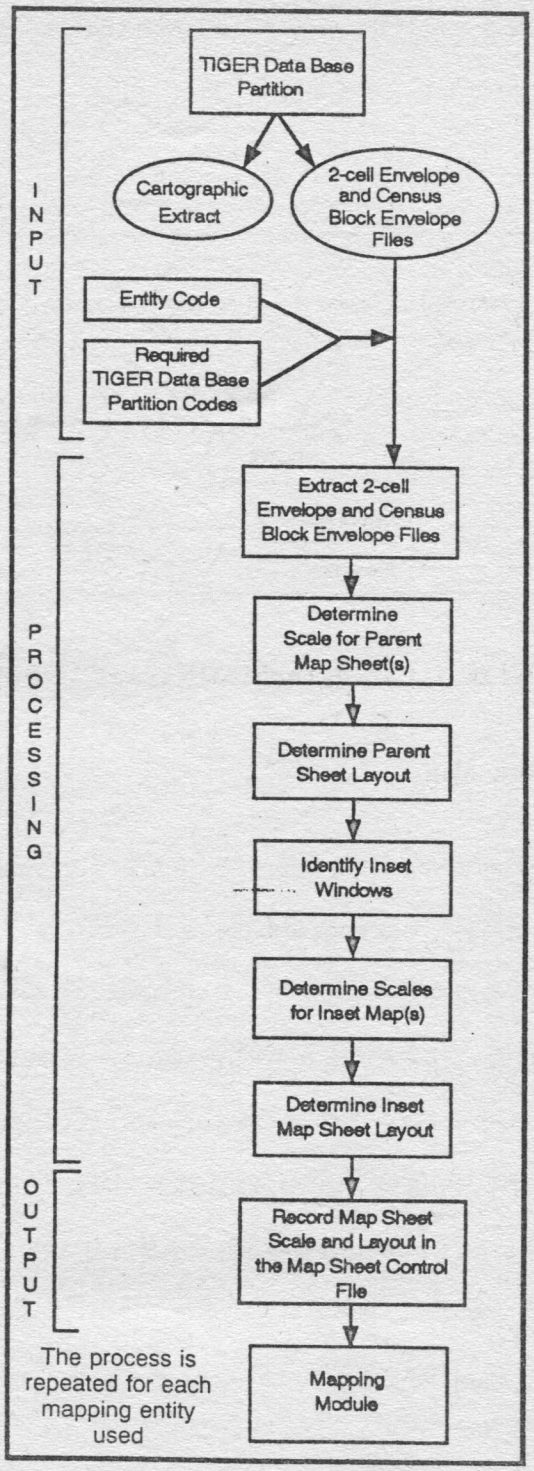


〔그림 22〕 **Cartographic Extract의 구조**

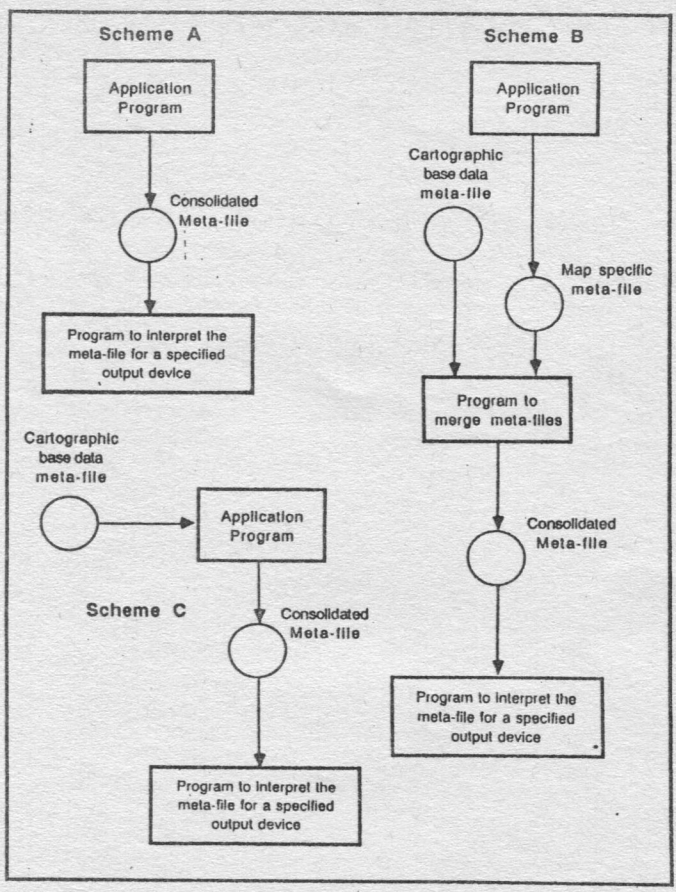
4.4 Publication Map

Cartographic extract가 조사구 요도를 제작하는데 필요하다면, 발간용 통계지도 작성에는 cartographic meta-file을 사용한다. 조사구 요도와는 달리 발간 통계지도는 종류가 작고 높은 질을 요구하기 때문에 interactive mode를 이용하며, non-interactive mode도 이용된다.

하나의 지도를 제작하기 위해서는 하나의 cartographic meta-file이 필요한데, 지도 제작에 필요한 meta-file에는 자료(object definition) 뿐만 아니라 function(subroutine)이 수록되어 있다. Cartographic extract와는 달리 축척·내용 등이 비슷한 다른 지도를 위해 제작된 meta-file을 가감하여 새로운 meta-file을 만들 수 있다.〔그림 24〕.



〔 그림 23 〕 축척, 도엽, 삽도결정의 자동화



〔그림 24〕 Meta-file과 통계지도의 제작

4.5 TIGER시스템의 성공 요건

지난 30년간의 GIS 역사에는 성공보다는 실패가 더 많았었다. 이가운데 TIGER 시스템은 GIS 기술축적이 가져온 개가라 할 수 있다. TIGER 시스템은 우리나라의 지역정보 전산화 또는 센서스 전산화 계획 및 수행에 많은 시사점을 주고 있다. 이 시스템을 성공으로 이끌었던 몇가지 요소를 정리하면 다음과 같다.

① Top management의 신기술에 대한 과감한 투자결단.

② GIS 기술의 성숙(hardware/software).

③ 조사통계국의 훌륭한 인적자원과 DIME file 제작을 통한 다년간의 축적된 경험.

④ 기존 직원(지도학자)의 경험과 신규채용 직원(프로그래머)의 신기술이 조화로운 결합을 이룸.

⑤ 국가 기본도가 준비되어 있었고, USGS가 DLG 등 수치지도 데이타를 제공하는 등 기관사이의 협조

〈부　　록〉

〔부록〕 미 조사통계국의 1990 센서스 조사용 지도

I. 1990 Census Field Maps

Field maps are designed to help with the collection of data for the 1990 decennial census. A map type may be used for a variety of decennial census field and office operations. To serve all requirements, map types often are modified to meet the needs of a specific operation. Field maps can be placed in three categories: 1) true field maps; 2) boundary collection/verification maps; and 3) early/limited use data review maps. True field maps are used only by Census Bureau staff for data-collection activities. Boundary collection/verification maps are used by Census Bureau staff and various local government and tribal officials for geographic entity boundary update and verification. Early/limited use data review maps are used by Census Bureau staff and various local government and tribal officials for additional geographic entity definition (census block equivalency) and for data review. These maps also are made available to the general public. The three categories of field maps are described below. (The highlighted numbers displayed parenthetically following each map type title show the number of map sheets being generated for the 1990 decennial census.)

A. True Field Maps

1. **Address Register Area (ARA) Map (311,835)**—A large-scale map produced for each ARA. Collectively, the groups for the four types of 1990 census enumeration cover the entire United States and its possessions: 1988 prelist; precanvass (TAR); 1989 prelist (update/leave); and list/enumerate. A limited number of geographic entity boundaries, including state, county, AI/ANA, and ARA, are shown. Areal features such as water bodies; linear features such as roads, railroads, pipelines, power transmission lines, and single-line hydrography; and landmark features such as mountain peaks, parks, and schools are displayed and named. Census block numbers are displayed. Listings identifying all streets and their associated ARAs within district offices may be used as a cross-reference to ARA Map sheets.
2. **Block Sketch Map (4,983)**—A large-scale map representing either a single census block or a cluster of census blocks. Features within the subject census block or census block cluster (including census block numbers) are identified. A list identifying all census block numbers in a specified number of surrounding tiers of blocks is included with each map for reference.
3. **Block Split Map (110,100)**—A large-scale map produced for each ARA. Cycle 1 maps include blocks split by geographic entity boundaries in selected TAR and 1988 prelist ARAs, while Cycle 2 maps include blocks split by such boundaries for all types of enumeration areas, except those ARAs previously mapped in Cycle 1 containing no new split blocks. In addition to the geographic entities shown on the ARA Map, other geographic entities shown on this version include MCD/CCD, incorporated place/CDP, UA, VTD, and CD. The boundaries of these tabulation areas, their names or identifying numbers, and tabulation block numbers are shown. Address range information is included in TAR areas only. Areal, linear, and landmark features are shown.
4. **County Locator Map (8,266)**—A medium-scale outline map used primarily by the DOs for planning, orientation, recruiting, and managing enumerator assignments. Map coverage includes all territory in each county. State, county, DO, AI/ANA, and ARA boundaries are shown. MCDs, incorporated places, and ANVSAs are identified by name only. All cartographic base features coincident with displayed boundaries are shown and major noncoincident features, such as roads and streams, are displayed. Cross-reference listings identifying the ARA numbers on a map sheet and the map sheets associated with a specific ARA number are provided with the map for reference.
5. **District Office Wall Map (4,827)**—Multiple full-size map sheets that can be mosaicked to display the service area of an entire DO. The Urban District Office Wall Map is a medium- to large-scale map produced for selected urban area DOs, while the Rural District Office Wall Map is a small-scale map designed for DOs covering a large geographic area, typically several counties. Within the DO, major features are shown along with geographic entities including county, CT/BNA, and ARA. Census block numbers are also shown on the Urban District Office Wall Maps.
6. **Postal Locator Map (1,404)**—A small-scale map produced for each mail out/mail back area located within an update/leave or list/enumerate county. State and county boundaries are shown. Selected cartographic base features are placed and named. The mail out/mail back area is highlighted by applying a screened shading over the adjacent fringe area.

B. Boundary Collection/Verification Maps

1. **Boundary and Annexation Survey Map (141,864)**—A map produced for each functioning county, MCD, and incorporated place. Maps are sent to local officials for update and certification of geographic entity boundaries. The 1988 BAS maps were designed to collect boundaries current as of January 1, 1988, whereas the 1990 BAS maps were designed to collect January 1, 1990, boundaries. Additionally, "BAS like" maps were produced for Puerto Rico, the Virgin Islands of the United States, and the Pacific Outlying Areas.
2. **Collection Insertion Map (74,063)**—Collection Insertion Maps were quad-based, fixed-scale maps created for editing insertion to the TIGER data base of 1990 census tracts and block numbering areas.
3. **Tribal Review Map (1,771)**—A medium-scale map produced for each AI/ANA and used by tribal officials for update and certification of boundaries.

C. Early/Limited Use Data Review Maps

1. **County Block Map (209,600)**—A medium-scale, county-based map, normally consisting of multiple sheets for each county. The maps portray the boundaries and names or identifying numbers of most geographic entities for which the Census Bureau tabulates data, as well as a wide range of cartographic base features. Census block numbers are shown.
2. **Governmental Unit Map (203,334)**—Governmental Unit Maps are produced for each MCD, incorporated place, and AI/ANA for which the Census Bureau tabulates data. They are medium-scale maps, identical in content and design to the County Block Map.
3. **TAR Geocoding Resolution/Work Place Spotting Map (62,123)**—A large-scale, census tract-based map used for researching unmatched addresses and identifying the locations of major employment sites. Geographic areas identified include state, county, AI/ANA, and CT/BNA. Areal features, linear features, and landmark features are shown. Address range information is included.

II. 1990 Census Data Product Maps

Data product maps are designed for use with the data the Census Bureau tabulates as a result of each census and sample survey. Most data product maps portray the boundaries of the political and statistical entities for which the Census Bureau tabulates data. Some also present the spatial distribution and relative magnitude of sets of census data. Data product maps can be placed in four categories: 1) data user maps; 2) printed reference outline maps; 3) printed thematic maps (page-size); and 4) printed thematic maps (GE-50 and GE-70 series). Both data user maps and printed reference outline maps are summary reference outline maps. Data user maps are electrostatically plotted, monochromatic, paper products. Printed reference outline maps are included in the published 1990 census reports and are printed in the traditional sense. Printed thematic maps are categorized according to size; page-size maps are approximately 8.5 x 11 inches while the GE-50 and GE-70 series maps are approximately 36 x 42 inches.

A. Data User Maps

1. **1990 County Block Map Series and Index (180,000)**—A comprehensive, multi-sheet, county-based reference map series. The maps portray and identify most geographic entities for which the Census Bureau tabulates data and a wide range of cartographic base features. Census block numbers are shown. An index to map sheets is provided for each county. The maps are prepared initially to accompany the special population summary prepared in accordance with P.L. 94-171. As such, they show the boundaries and identifying numbers of all voting districts submitted by participating state officials (90,000 map sheets). A second version of the map, on which voting district boundaries are omitted, is released to accompany the printed 1990 census reports and STFs (90,000 map sheets).
2. **Census Tract/Block Numbering Area Outline Map (4,200)**—A full-size (36 x 42 inches), county-based map series. Identifies MCD/CCDs, incorporated places/CDPs, AI/ANAs, and CT/BNAs. Only those cartographic base features coincident with CT/BNA boundaries are identified.
3. **County Subdivision Outline Map and Index (100)**—A state-based map series, generally available as a single, full-size map sheet. Identifies counties, MCD/CCDs, incorporated places/CDPs, and AI/ANAs.
4. **Voting District Outline Map (4,000)**—A full-size, county-based map series, produced for counties participating in Phase 2 of the 1990 Census Redistricting Data Program. Identifies MCD/CCDs, incorporated places/CDPs, and VTDs. Distinguishes between types of VTDs (actual and pseudo). All cartographic base features coincident with VTD boundaries are identified.
5. **Urbanized Area Outline Map, with selected base features (400)**—A full-size, UA-based map series with state, county, MCD/CCD, incorporated place/CDP, AI/ANA, and UA boundaries and names identified and major roads shown. All cartographic base features coincident with the UA boundary are symbolized and named. A fringe area is included and highlighted to distinguish the extent of the UA.

B. Printed Reference Outline Maps

1. **American Indian/Alaska Native Areas Outline Map (500)**—A page-size, AI/ANA-based map series. State, county, MCD/CCD, and incorporated place/CDP boundaries and names are shown in addition to the AI/ANA.
2. **American Indian/Alaska Native Areas of the United States (1)**—A two-page-size map of the United States showing state, county, and AI/ANA boundaries and names.
3. **Census Tract/Block Numbering Area Outline Map (4,200)**—A full-size (36 x 42 inches), county-based map series. The higher-quality version of the Census Tract/Block Numbering Area Outline Map available as a "data user" map.
4. **County Subdivision Outline Map and Index (300)**—A state-based map series, produced as a multi-sheet, page-size map series. Identifies counties, MCD/CCDs, incorporated places/CDPs, and AI/ANAs. An index to the map sheets is provided for each state.
5. **Congressional District Outline Map (550)**—A state-based, page-size map series bound in several Census Bureau publications and presented in the *Congressional District Atlas*. The boundary of each CD and the included counties are symbolized and identified. Selected incorporated places are shown. County-based inset maps are provided for counties containing two or more CDs. Cartographic base features coincident with CD boundaries are identified.
6. **Metropolitan Areas of the United States (1)**—A two-page-size map of the United States showing state and metropolitan area boundaries and names. Type style and boundary symbology distinguishes the three types of metropolitan areas—MSAs, CMSAs, and PMSAs.
7. **Public Use Micro-Data Sample (PUMS) Areas (18)**—Seventeen slightly overlapping sections covering the United States showing state, county, and PUMS "area" boundaries and names. An index to map sheets is produced.
8. **Regions and Divisions of the United States (1)**—A page-size map of the United States showing state, region, and division boundaries and names.
9. **State/County Outline Map (57)**—A page-size, state-based map series showing state and county boundaries and names. State capitals and selected places are included.
10. **State/Metropolitan Area Outline Map (57)**—A page-size, state-based map series showing state, county, and metropolitan area boundaries and names. The three types of metropolitan areas—MSAs, CMSAs, and PMSAs—are distinguished. State capitals and selected places are included.
11. **Urbanized Area Outline Map (400)**—A page-size, UA-based map series. State, county, MCD/CCD, incorporated place/CDP, and AI/ANA boundaries and names are shown in addition to the extent of the subject UA.
12. **Urbanized Areas of the United States (1)**—A two-page-size map of the United States showing state, county, and UA boundaries and names.

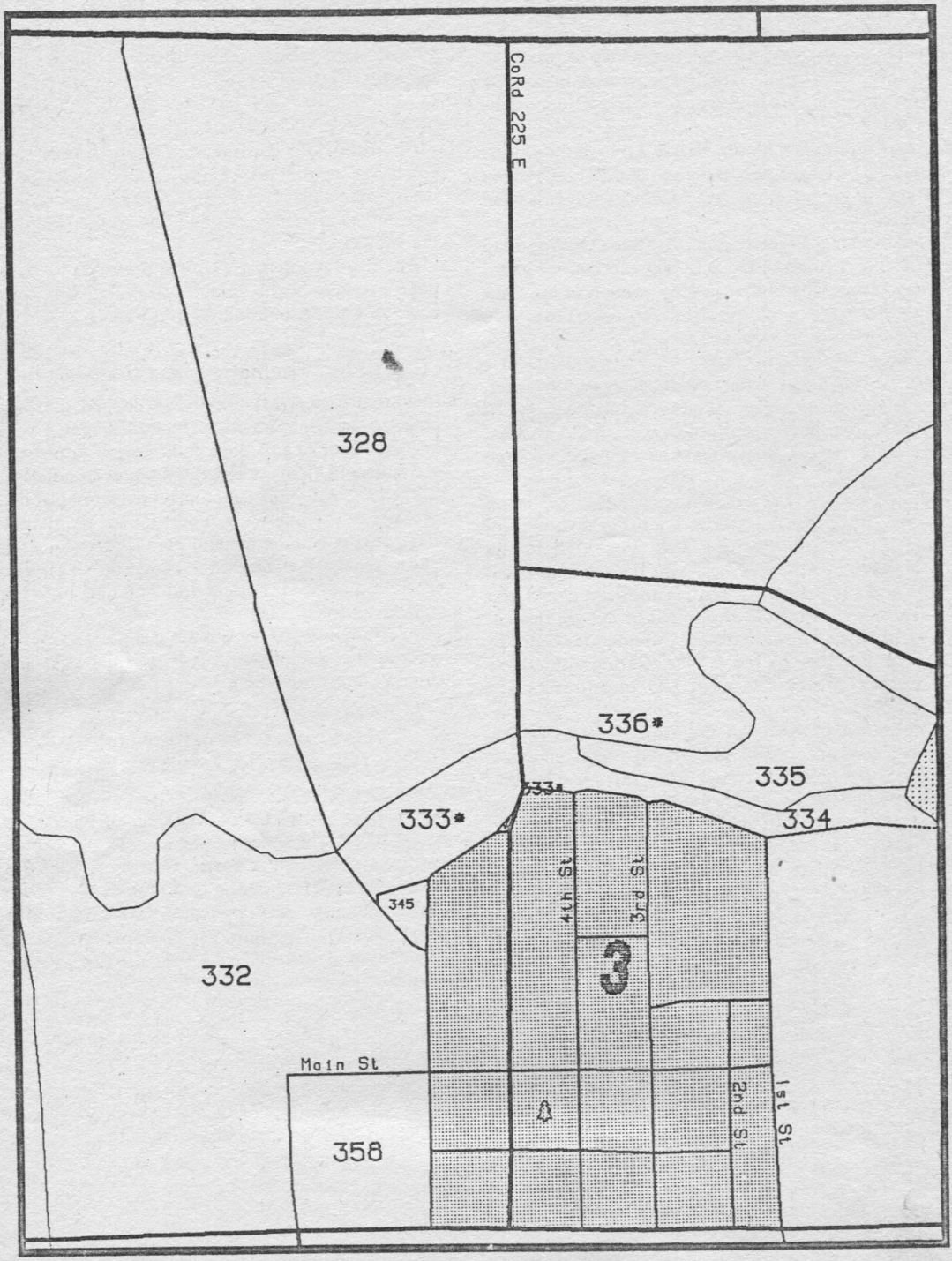
C. Printed Thematic Maps (Page-size)

1. **Assorted Titles (24)**—Page-size printed thematic maps of the United States. The two standard titles include Major Acquisitions and Dates of Admission of States and Centers of United States Population: 1790—1990. Additional choropleth maps are planned for each of the following report topics: population and housing counts; general population characteristics; general housing characteristics; social and economic characteristics; and detailed housing characteristics.
2. **Population and Housing Atlas (580)**—A proposed volume that would contain numerous page-size United States thematic maps.

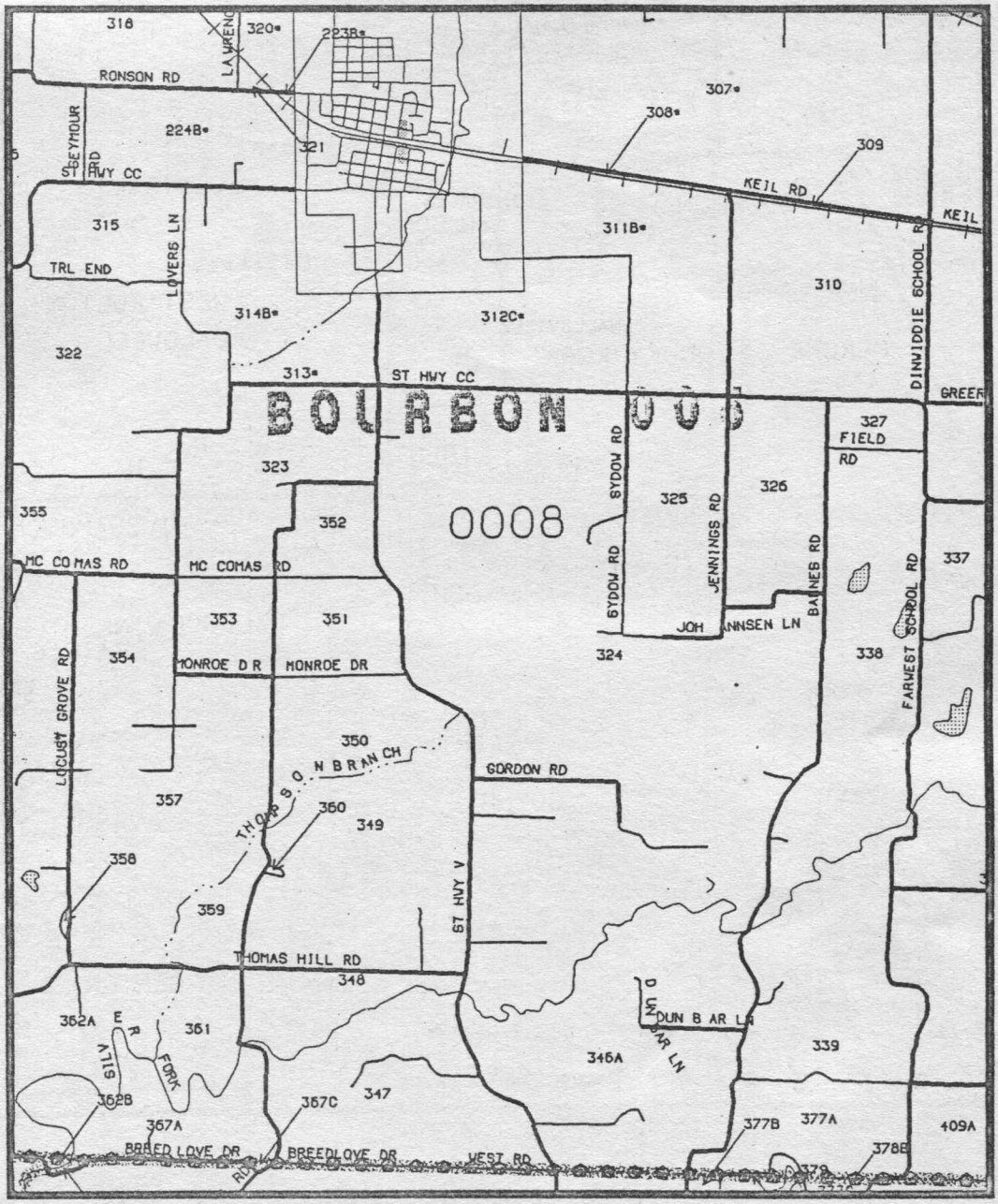
D. Printed Thematic Maps (GE-50 and GE-70 Series)

GE-50 and GE-70 Map Series (20)—Full-size, color maps of the United States, at a scale of 1:5,000,000 or 1:7,500,000. Standard titles include: Metropolitan Areas of the United States, Districts of the 103rd Congress, 1990 Population Distribution Map (Daytime View), and 1990 Population Distribution Map (Nighttime View). Additional choropleth titles are planned to depict selected 1990 census data by county.

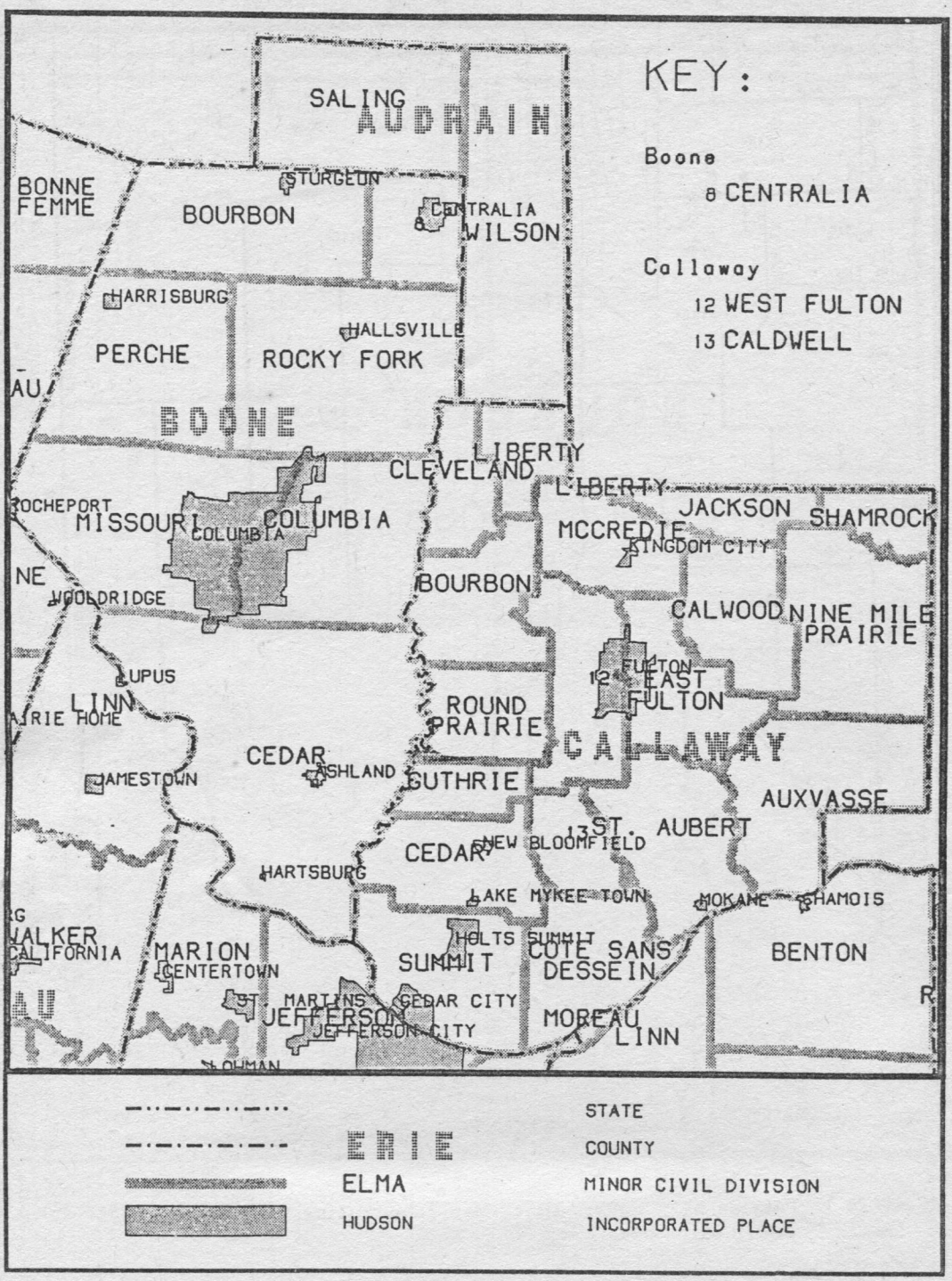
Additional information regarding the wide range of Census Bureau map products may be obtained by contacting Customer Services Branch, Data User Services Division, Bureau of the Census, Washington, DC 20233. Phone: 301/763-4100.



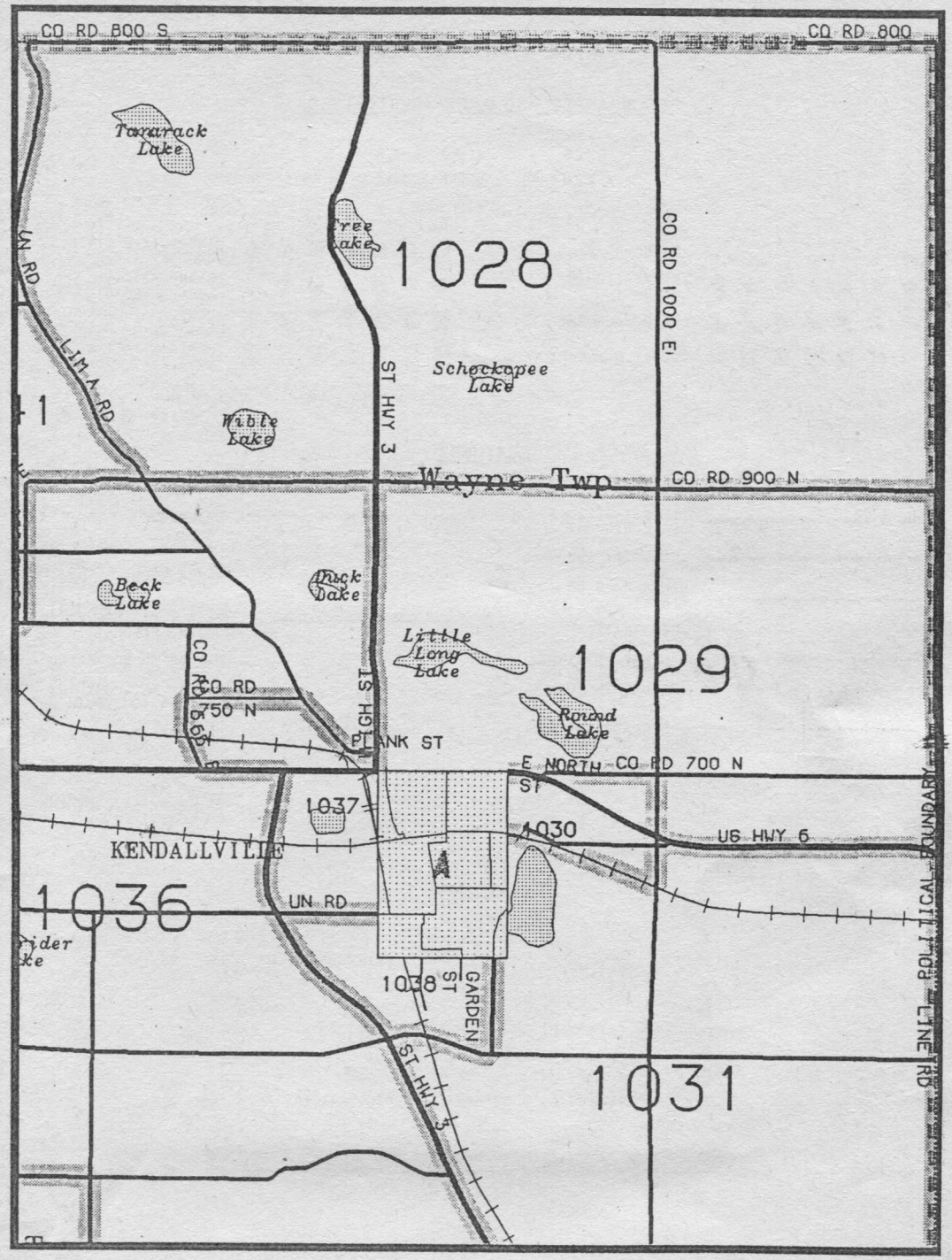
〔그림 25〕 **Portion of an Address Register Area illustrating an area for whieh the mapping software produced an inset by whole blocks.**



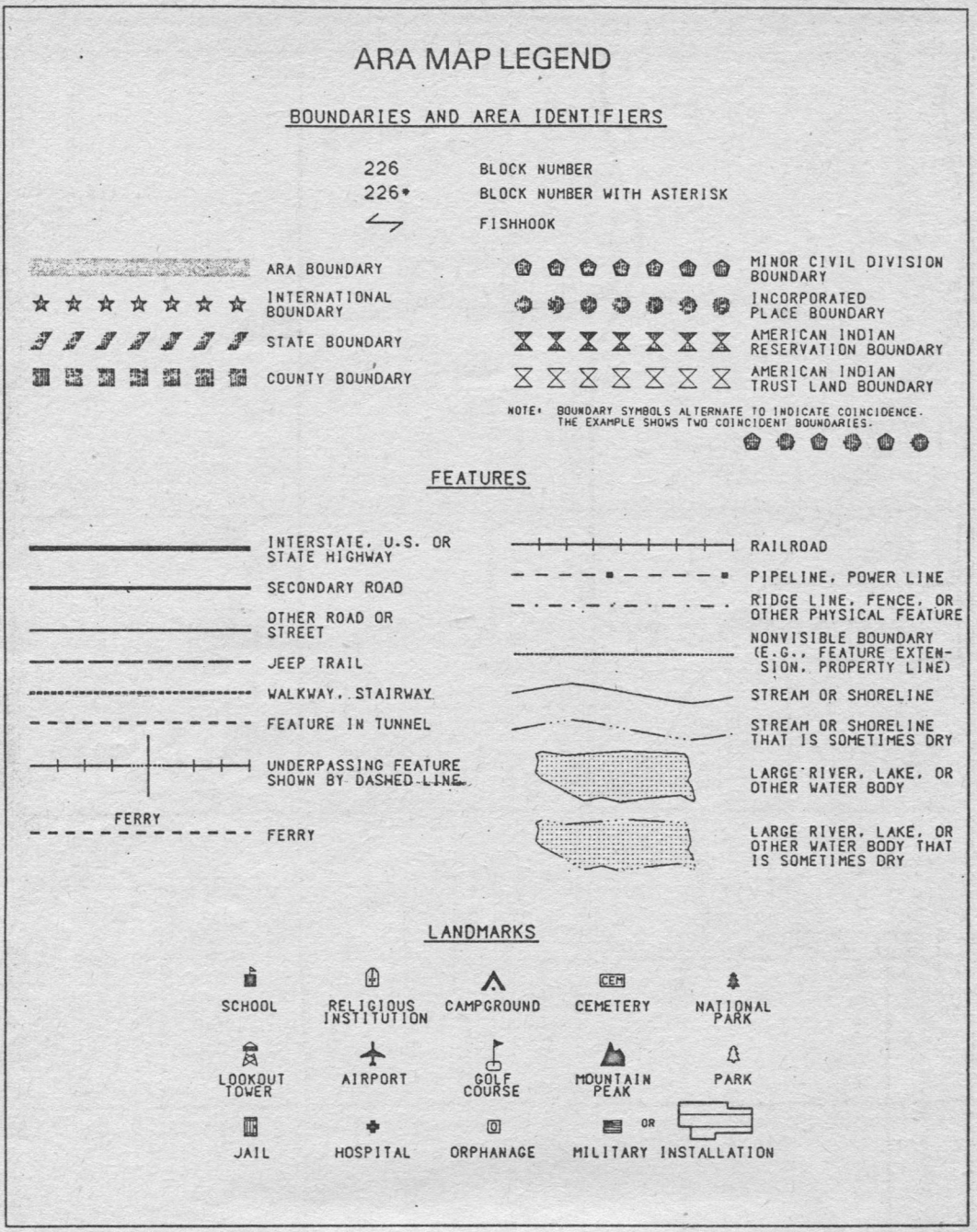
〔그림 26〕 **Portion of a County Block Map illustrating rectilinear insets for small-scale maps.**



〔그림 27〕 **Portion of a County Subdivision Map illustrating areal feature names Placement for governmental units and the use of key numbers.**



〔그림 28〕 **Portion of a County Locator Map illustrating the suppression of text within an inset area.**



〔그림 29〕 **Example of an Address Register Area Map legend.**