

# 지리정보를 갖는 통계 데이터의 Visualization

이정진<sup>1)</sup>

요 약

정보화시대의 발전과 더불어 우리 일상생활에 친숙하게 다가온 기술 중의 하나가 지리정보시스템(Geographical Information System: GIS)이다. GIS는 공간(지표, 지하, 해양 등)상에 분포하고 있는 정보에 대해 여러 종류의 세밀한 지도를 이용하여 효율적으로 사용자에게 제공하여 관리하는 종합정보기술이다. 통계 데이터 중에는 지리정보를 가지고 있는 경우가 상당히 많다. 하지만 지리정보를 갖는 통계 데이터의 탐색적 자료분석(Exploratory Data Analysis)을 위한 Visualization 기법에 대해서는 별로 연구된 바가 없다. 본 논문에서는 GIS를 위한 벡터맵(vector map)의 간단한 제작방법과 GIS의 개발방법, 그리고 탐색적 자료분석을 위한 Visualization기법을 소개한다.

주요용어 : Geographical Information System, Visualization

## 1. 통계 데이터를 위한 GIS의 개발

상용되는 GIS에 사용되는 지도는 대부분 위성 사진 등을 이용한 매우 정교한 것으로 비용과 시간이 많이 요구되어 대부분의 GIS 시스템은 가격이 매우 비싸다. 하지만 통계데이터를 위해서는 이러한 고가의 정교한 지도가 필요 없고, 우리나라의 경우 대략의 시/도별 또는 시/군/구별의 경계를 구분할 수 있을 정도면 충분하다. 컴퓨터시스템을 위한 지도의 제작방법은 여러 가지가 있으나 최근에는 얼마든지 확대 축소 등의 브라우즈(browse)가 가능한 벡터맵(vector map)을 이용한다. 벡터맵이란 한 지역의 모양을 잘 나타낼 수 있는 X축 Y축좌표를 촘촘히 구하여 이들 좌표들을 직선 또는 곡선으로 연결함으로써 지도가 나타나도록 하는 것이다. 상용으로 판매도 되고 있으나 본 시스템을 위해서는 정교한 지도는 필요치 않다고 판단되어 시범적으로 직접 좌표를 추출하여 벡터맵을 제작하였다

지리정보를 갖는 통계데이터를 표현하기 위해서는 자료의 입력, 사용자가 원하는 visualization을 위한 대화상자의 구축, 그리고 벡터맵을 적절히 운용하는 모듈(browse, drill down action 등), 그리고 visualization된 결과를 출력하는 Geographic Information System의 구축이 필요하다. 본 시스템에서는 기존의 S-Link 데이터 입력 시스템과 결과출력시스템을 이용하였고 운용 모듈은 Visual C++를 이용하여 직접 제작하였다.

## 2. 지리정보를 갖는 통계데이터의 Visualization

### 2.1 이산형 조건변수를 갖는 데이터의 GIS Visualization

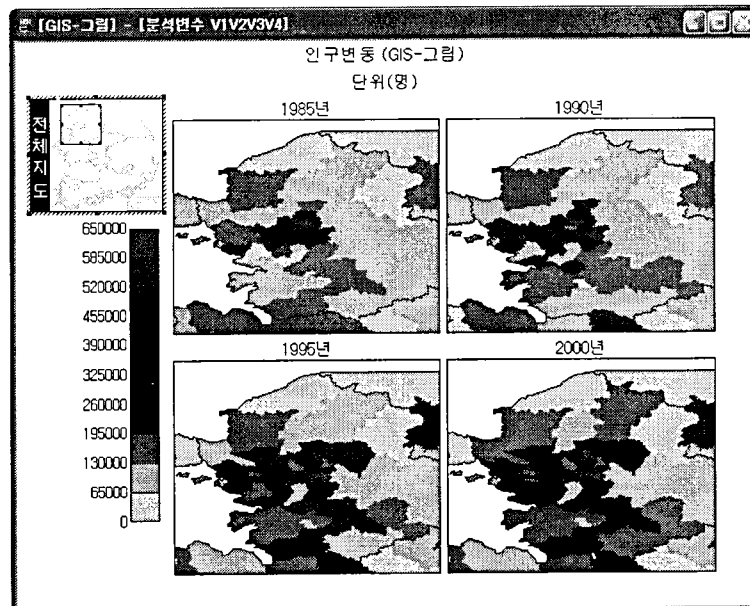
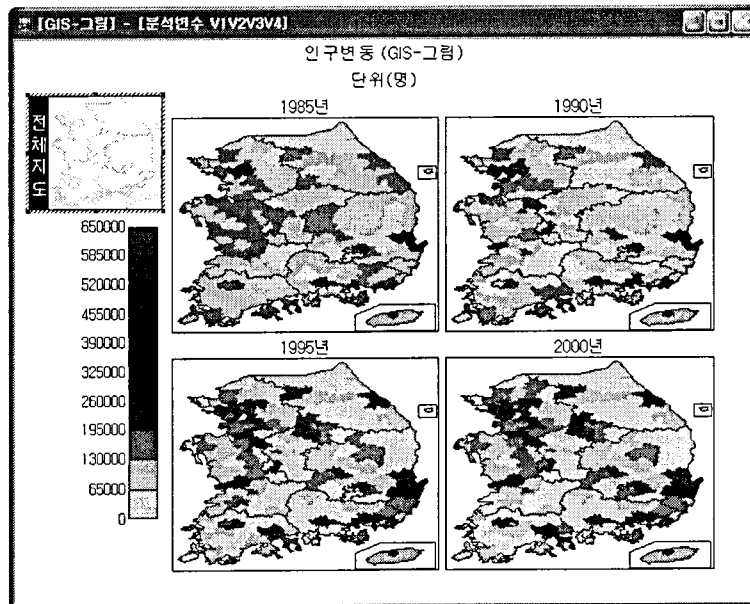
<그림 1>의 왼쪽은 1985, 1990, 1995, 2000년도 센서스 자료를 이용하여 우리 나라의 총인구를 각 시군구별로 구분하여 Visualization 한 것이다. 왼쪽 상단에 있는 지도를 마우스로 특정지

---

1) 156-743 서울시 동작구 상도동 1-1, 숭실대학교 정보통계학과 교수, E-mail: jjlee@ssu.ac.kr

지리정보를 갖는 통계 데이터의 Visualization

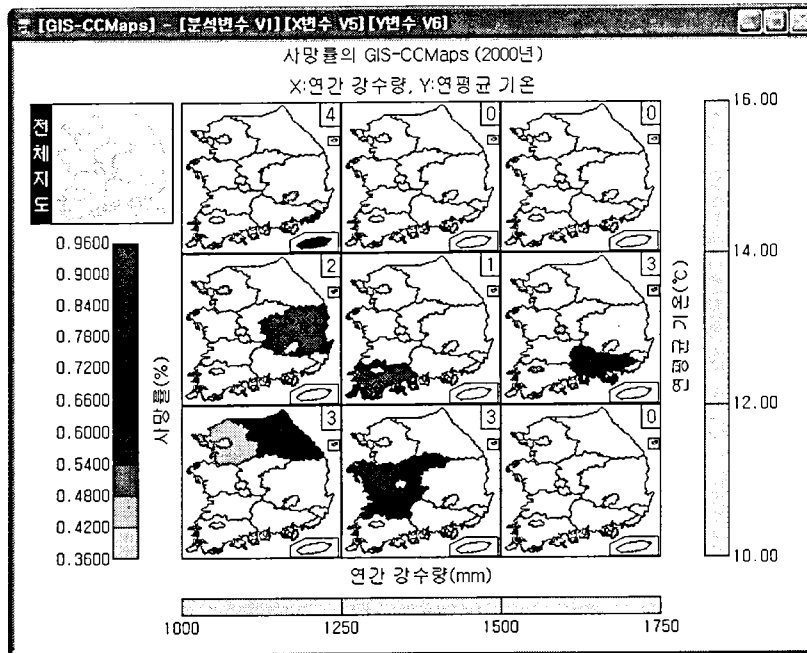
역을 선택하면 해당되는 지역에 대해 확대를 할 수 있는데, 수도권에 대해 확대한 그림이 오른쪽 그림이다. 수도권과 부산·경남 지역의 인구가 많이 증가되었음을 쉽게 알 수 있고 증기능을 활용하면 구체적인 해당 시군구를 알아 낼 수 있다. 왼쪽에 인구수의 구분을 나타내는 바(bar)의 간격은 사용자가 마우스를 이용하여 또는 원하는 대로 조정할 수도 있다.



<그림 1> 각 시군구별 인구의 연도별 visualization 및 수도권에 대해 확대한 그림

2.2 연속형 조건변수를 갖는 데이터의 GIS Visualization

<그림 2>는 우리 나라 각 시군구의 인구에 대비한 사망률을 해당 지역의 연 평균기온과 연간 강수량을 조건변수로 하여 visualization한 그림이다. X축과 Y축은 디폴트로 3개의 구간으로 나누게 되어 있고 해당 톨바의 경계선을 마우스로 이동하면서 지역별 변화를 관찰할 수 있다. 왼쪽 상단에 있는 지도를 마우스로 특정지역을 선택하면 해당되는 지역에 대해 확대를 할 수 있다. 대구·경북지역과 전남 지역의 사망률이 매우 높음을 알 수 있고, 이 지역은 연 평균기온이 12도에서 14도 사이이고 연간 강수량은 1500mm 이하이다. 역시 왼쪽에 인구수의 구분을 나타내는 바(bar)와 의 간격은 사용자가 마우스를 이용하여 원하는 대로 조정할 수 있어 세밀한 지역의 사망율을 잘 관찰할 수 있다.



<그림 2> 각 시군구의 인구에 대비한 사망률을 해당 지역의 연평균기온과 연간 강수량을 조건변수로 하여 visualization한 그림

3. 토의

본 논문에서 소개한 통계 데이터를 위한 GIS의 개발과 Visualization software는 시범 제작한 prototype으로서 향후 실용화를 위해서는 많은 수정 및 발전이 필요하다고 생각된다. 관심 있는 많은 분들의 조언을 바란다.

참고문헌

오종우 (2000), 지리정보시스템, 현대기술정보(주)