

구조적인 관계를 이용한 지도의 정보추출에 관한 연구

전흥구, 최관순
순천향대학교 정보기술공학부

A Study of Information Extraction from Map using a Structural Relation

Heung-Goo Jun, Kwan Sun Choi
School of Information and Technology, Soonchunhyang Univ*

요 약

지도영상은 숫자, 각종 심볼 그리고 지형적인 경계선, 도로, 강 등을 나타내는 라인 등으로 구성되어 있다. 본 연구에서는 기하학적인 관계를 이용하여 지도영상으로부터 지도정보를 추출하는 연구를 수행하였다. 제안된 이 방법은 연결성분만을 고려한 방법보다 개선된 방법으로 GIS 입력 모듈에 효율적으로 이용할 수 있다.

abstract

A typical map image is consisted of number, symbols and lines that represent contour lines, geographical boundaries, roads, rivers and etc. This paper describe the separation. The separation algorithm is based on structural relation. The separation of proposed method is better than method using only a size of connected component. This proposed method will be used in GIS input system efficiently.

1. 서 론

일반적으로 지도는 이용 목적이 다양하고, 수요가 많은 일반도와 전문적인 용도에 따라서 제작된 특수용도의 주제도로 구분된다. 지도에는 등고선, 해안선, 호수, 하천 등 지형정보와 도로, 철도, 건축물 등 지리정보가 기록되어 있으며 각종 기호로 나타내고 있다. 일반적으로 자연지형이나 인공구조물 등은 3차원 공간에 존재하지만 2차원 평면으로 표현되어야하므로 지형은 3차원의 형상이 2차원 평면상에 등고선과 표고값 등으로 표현된다. 또한 점, 선, 면으로 구성되어진 지리정보가 용이하게 분별될 수 있도록 점으로 표현되는 인공물을 지도위에 기호로 나타내며, 도로는 선으로, 경계선 등은 선의 종류 즉 실선, 파선, 쇄선 등과 선의 굵기로, 면으로 표현되는 건축물 등은 색상과 패턴으로 표현된다. 수치지도는 컴퓨터의 영상처리기술을 이용하여 이러한 정보들을 레이

어로 구분하여 데이터베이스화하고 이용목적에 따라 지도를 자유로이 작성할 수 있도록 전산화한 지도로 종이지도를 전산화하는 과정을 의미하며 이 수치지도는 GIS구축에 바탕이 된다. 신규 지형도의 제작에는 많은 경비가 소요된다. 예를 들면, 일본의 경우 시설관리시스템에서 표준적으로 사용되는 축척 500분의 1 지형도에서는 시가지에서 1km당 약 600만원의 비용이 필요하다고 한다. 따라서, 지형도를 일단 제작한 지방 자치단체에서는 지형도의 갱신 시에는 특별한 사정이 없는 한, 기존 지형도의 변화한 부분만을 수정하여 사용하고 있는 실정이다.

수치지도 데이터 입력 시스템으로서 현재 3가지가 이용되고 있다.

- 디지털 입력 시스템
- 대화형 (반자동) 입력시스템
- 자동입력 시스템

이 중에 디지털 입력 시스템이 현재 가장 많이 사용되고 있지만 인간의 손에 의해 지형도의 한 점 한 점을 입력하는 시스템이기 때문에 입력 효율이 대폭적으로 개선될 가능성이 없다.

대화형 입력 시스템은 인간에게 도형인식을 맡기고 벡터화만을 자동적으로 하는 것으로 이 시스템에서는 기존 지형도의 영상이 칼라 화면에 표시되고 인간이 도형의 의미 (레이어; 도로, 가교, 건물 등)를 지정한 다음에 벡터화할 도형을 지정한다.

자동 입력 시스템은 스캐너를 이용하여 기존 지형도를 라스터 영상으로 변환한 후 벡터화하고 도형을 인식하여 레이어 분할을 한다.

2. 본 론

본 연구에서는 기존의 종이지도를 문서로 보관 하는 것보다 이를 수치화하여 데이터베이스화 하게 되면 많은 응용부분에서 가공하여 활용할 수 있기 때문에 기존의 지도로부터 정보를 추출하기 위해 구조적인 관계를 이용한 방법을 제안하였다. 제안한 방법은 다음과 같은 과정을 거친다.

1) 지도영상의 취득

이미지스캐너를 사용하여 2진 지도영상을 해상도가 300dpi로 영상을 취득하였다. 2진 영상에서 가장 작은 단위는 화소로 같은 값을 갖는 화소가 서로 그룹을 지어 형태를 이루며 심볼이나 숫자, 등고선, 하천 등과 같은 도형 및 비도형을 나타낸다.

가. 짧은 라인과 긴라인의 분리

지도 영상에 나타나는 심볼과 숫자는 화소수와 영역을 둘러싸는 최소사각형이 작기 때문에 짧은 라인에 해당하고 등고선이나 도로, 건물 라인은 긴라인에 해당되며 이는 영역확장 알고리즘을 사용하여 분리한다.

나. 노이즈 성분

연결 요소의 가로, 세로의 길이가 임계값(2화소)이내에 들면 먼저 노이즈로 간주하며, 이것들 중에 숫자들 사이에 존재하는 노이즈 성분은 소수점에 해당하므로 숫자열로 합류시킨다.

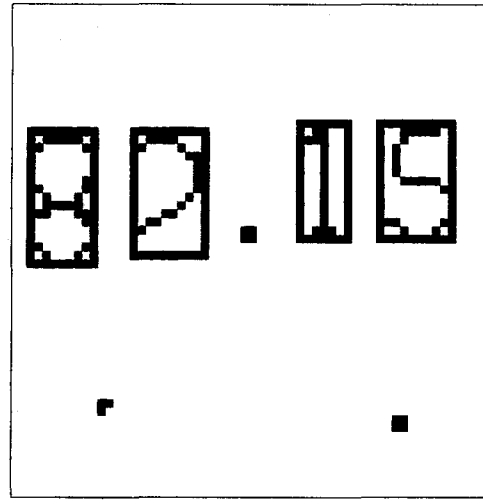


그림 1. 소수점과 잡음

다. 사선(Dashed line)의 분류

작은 연결 요소들로부터 다음과 같은 사선 연결 요소들이 갖는 특성을 이용하여 숫자, 심볼 들보다 먼저 추출한다. 그림 5에서 연결 요소를 이루는 최소사각형의 X축, Y축으로 투영하였을 때 어느 한쪽은 균일한 분포를 이루며, 흑백화소의 변화도 2이다.

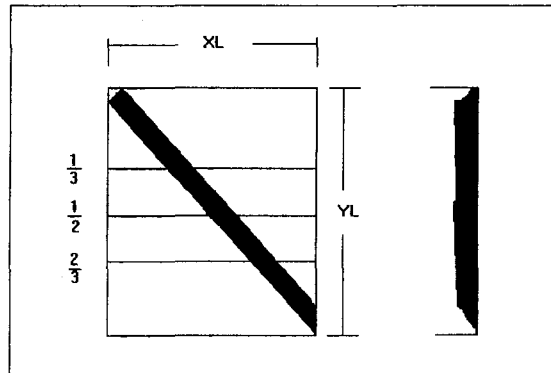


그림 2. 사선패턴의 히스토그램과 화소 변화율

라. 연결 요소의 그룹핑

심볼 요소 중 우체국, 도근점, 삼각점, 발, 풀밭은 2개 이상의 연결 요소로 이루어져 있으므로 하나의 심볼로 인식하기 위해 연결 요소의 그룹핑이 요구된다.

(1) 포함 : 그림 3과 같이 큰 연결 요소가 작은 연결 요소를 포함한다.

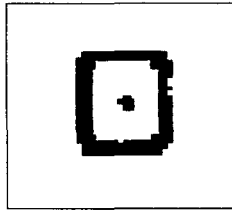


그림 3. 도근점

(2) 겹침 : 그림 5와 같이 연결 요소를 둘러싸는 최소 사각형의 일부가 겹친다.

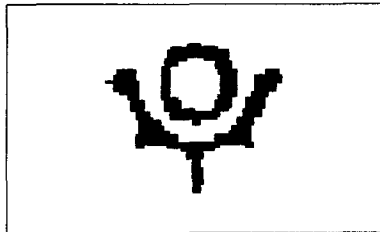


그림 4. 우체국

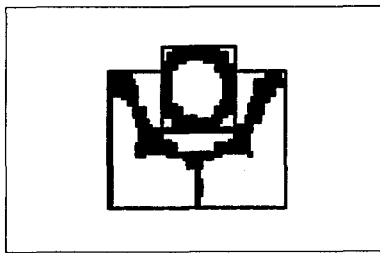


그림 5. 최소사각형으로 둘러싸인 우체국

(3) 평행 : 그림 6과 같이 2개 혹은 3개의 사선으로 이루어지며 사선끼리는 평행하다.

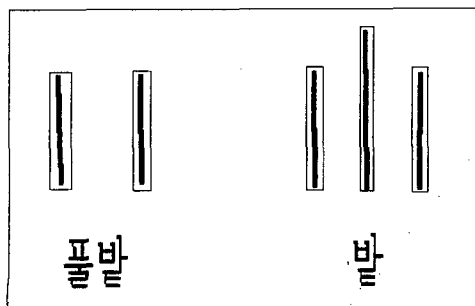


그림 6. 사선으로 이루어진 풀밭과 밭 기호

3. 시뮬레이션 결과

1) 연결 요소에 의한 심볼 및 숫자 추출 결과
 등압선도와 축적 1 : 5000 지도를 대상으로 스캐너를 사용하여 300 dpi 해상도로 이진화 지도 영상으로 그림 7과 그림 8은 등압선도이며, 그림 9는 지도영상으로 영상 (a)로부터 심볼 및 숫자를 추출한 결과를 (b)로 나타내었다.

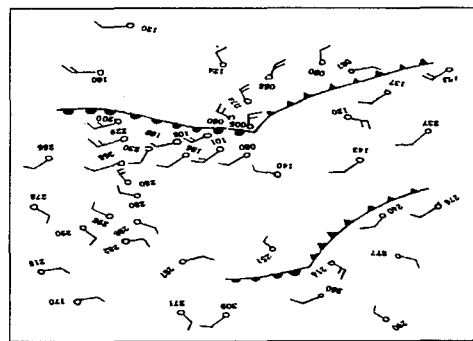


그림 7. (a) 등압선 영상 1

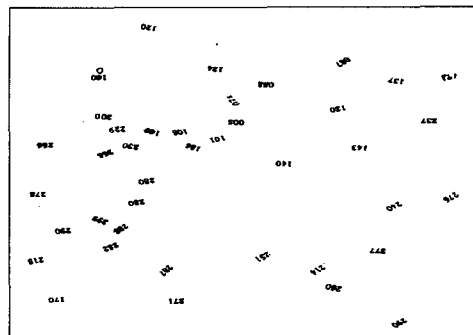


그림 7. (b) 등압선 영상 1에서 숫자열을 추출한 영상

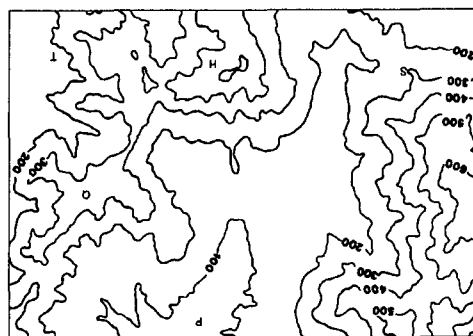


그림 8. (a) 등압선 영상 2

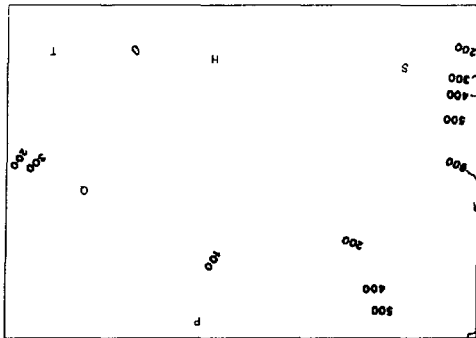


그림 8. (b) 등압선 영상 2에서 숫자열을 추출한 영상

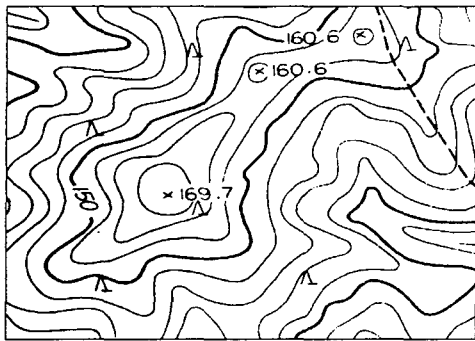


그림 9. (a) 테스트 지도영상 1

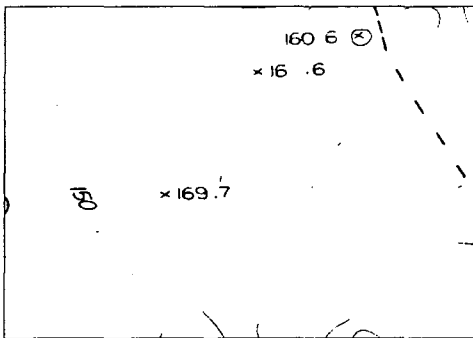


그림 9. (b) 테스트 지도영상 1에서 숫자열을 추출한 영상

4. 결 론

본 연구에서는 기존의 도면을 스캐너를 사용하여 래스터영상을 얻은 후 영상처리방법을 이용하여 구조적인 관계를 이용한 방법을 이용하여 지도상에 포함된 각종 데이터를 추출하는 연구를 수행하였다. 우선 연결성을 이용하여 긴 라인과 짧은 라인으로 분리한다. 짧은 라인은 보통 심볼,

숫자, 문자, 잡음 등이 이에 해당되며, 긴 라인은 도로, 하천, 등고선, 건물, 철도 등이 이에 해당된다. 짧은 라인에 속하는 이미지상에 나타나는 각종 심볼들은 기하학적인 관계를 이용하여 분리된다. 각종 라인들은 세션화과정을 거쳐 분리되어 표준적인 벡터양식 DXF 양식으로 저장된다. 제안 방법은 지도 이미지로부터 자동적인 정보추출과 이해를 위한 부분에 유용하게 이용할 수 있다.

참고문헌

1. H.samet *et al.*, "A geographic information system using quadtrees," *Pattern Recognition*, vol. 17, pp. 647-656, 1984.
2. M.Tavakoli *et al.*, "Building and road extraction from aerial photographs," *IEEE Trans. SMC.*, vol. 12, pp. 84-90, 1982.
3. T. J. Amin and R. Kasturi, "Map data processing - Recognition of lines and symbols," *OPT. ENG.*, vol. 26, pp. 354-358, 1987.
4. T. Ougaki *et al.*, "Automatic character extraction for urban maps," The 29th Annual Co nvention *IPSJ*, 6M-7, 1984 (in Jpanese).
5. M. Eijiri *et al.*, "Automatic recognition of designs ans maps," *proc. 7th ICPR*, pp.1296-1305, 1984.
6. R. Kasturi *et al.*, "A system for Interpretation of Line Drawings," *IEEE trans. Pattern Analysis and Machine Intelligence*, Vol.12, No.10, 1990, pp.978-991.
7. M.T. Musavi, *et al.*, "A Vision Based Method to Automate Map Processing," *pattern Recognition*, vol. 21, No.4, 1988, pp. 319-326.
8. R. Kasturi and J. Alemany., "InformationExtraction from Images of Paper-Based Maps," *IEEE Software Engineering*, Vol. 14. No. 5, May 1988.
9. M. Eijiri *et al.*, "Automatic Recognition of Engineering Drawings and Maps" in *Image Analysis Applicationa*, Marcel Dekker, pp 73-126, 1990.