The Journal of GIS Association of Korea, Vol.6, No.1, pp.11-23, April 1998

점패턴분석을 이용한 수치지형도의 점사상 일반화

유근배*

Generalization of Point Feature in Digital Map through Point Pattern Analysis

Keun-Bae Yu*

요 약

GIS 분야에서 지도 일반화는 공간자료의 상세도를 결정하여 효과적으로 자료를 가시화(Visualization)하거나 자료의 해상력을 변화시켜 축척을 변환하는 기능을 수행한다. 최근까지 지도 일반화는 선사상(Line Features)에 집중되었고, 수치지도를 구성하고 있는 정보량과 그 중요성에 비하여 점사상(Point Features)에 대한 연구는 상대적으로 미미하였다. 이러한 맥락에서 본 연구는 점사상에 대한 구체적인 일반화 방안을 모색하는데 목적을 둔다. 특히 점사상의 일반화에서 원자료의 기하학적 특성을 파악하는데 가장 중요하게 고려한 요소로 '점사상의 분포패턴'을 선정하였다. 즉 'Grieg-Smith 방법'을 활용한 방격분석(Quadrat Analysis)과 최근린분석(Nearest-Neighbour Analysis)를 통해 점사상이 갖고 있는 분포패턴의 특성을 찾아낸 다음, 이를 변형시키지 않도록 일반화의 기준거리(Threshold)를 설정하여 점사상을 제거하는 방법을 통해 점사상의 일반화를 시도하였다. 따라서 이 연구에서 제시한 점사상의 일반화 방안은 원래 점사상이 갖고 있는 기하학적 특성을 최대한 유지한다.

ABSTRACT : Map generalization functions to visualize the spatial data or to change their scale by changing the level of details of data. Until recently, the studies on map generalization have concentrated more on line features than on point features. However, point features are one of the essential components of digital maps and cannot be ignored because of the great amount of information they carry. This study, therefore, aimed to find out a detailed procedure of point features' generalization. Particularly, this work chose the distribution pattern of point features as the most important factor in the point generalization in investigating the geometric characteristics of source data. First, it attempted to find out the characteristics of distribution pattern of point features through quadrat analysis with Grieg-Smith method and nearest-neighbour analysis. It then generalized point features through the generalization threshold which did not alter the characteristics of distribution pattern and the removal of redudant point feautres. Therefore, the generalization procedure of point features provided by this work maintained the geometric characteristics as much as possible.

---

* 서울대학교 지리학과(Department of Geography, College of Social Sciences, Seoul National University, San 56-1 Shillin-dong, Kwanak-gu, Seoul 151-742, Korea)

1. 서 론

GIS는 다양한 분야에서 방대한 공간자료의 관리와 분석에 이용되고 있다. GIS의 기능을 충분히 발휘하기 위해서는 공간자료의 데이타베이스가 신속, 정확히 구축되어야 한다. 공간 데이터베이스를 구성하는 가장 기본적인 자료가 수치지형도이다. GIS의 사용자는 이용목적에 따라 다른 축척의 수치지도 데이터베이스를 요구한다. 그러나 사용자의 요구에 따라 같은 지역의 수치지형도를 다른 축척의 데이터베이스로 각각 구축한다면 자원의 손실과 중복투자를 초래하게 된다. 따라서 하나의 대축척 자료로부터 여러 가지 축척의 수치지형도로 변화하는 과정이 필요하다.

수치지형도는 크게 선형사상과 점사상으로 구성되어 있다. 지금까지의 수치지형도 일반화에 대한 연구는 선형사상에 집중되어 왔다. 선형사상의 정보양을 줄이면서도 원래의 선형사상의 특성, 즉 기하학적인 모양을 유지하는데 대부분의 연구가 진행되어 왔다. 그러나 점사상도 선형사상 못지 않게 수치지형도를 구성하는 주요 구성요소중의 하나이다. 점사상을 일반화하는 방법에 대한 연구는 선형사상 못지않은 중요한 의미를 지니지만, 점사상의 일반화에 대한 연구는 극히 미미한 실정이다.

점사상은 주위의 점사상들과 함께 특정한 분포 패턴을 형성하고 있다. 이러한 분포패턴은 점사상이 공간적으로 지니고 있는 특성을 나타낸다. 점사상의 일반화 과정에서는 분포패턴에 유의하여 이를 변화시키지 않도록 신중한 검토가 필요하다.

이러한 맥락에서 본 연구에서는 수치지형도의 주요 요소중 하나인 점사상을 축척별로 일반화하는 방안을 모색하도록 한다. 구체적으로 점사상의 일반화 방법을 살펴보고, 점분포의 패턴분석을 위한 기법을 모색한다. 마지막으로 사례지역을 선정하여 점분포패턴을 분석하고, 이를 유지할 수 있는 일반화 방안을 제시한다.

2. 점사상의 정의와 특성

점사상은 하나의 XY 좌표로 위치가 정의되는 지리적 사상이다. 점사상은 실질적인 공간차원을 가지고 있지 않는 지점 혹은 장소를 나타내는데, 단순한 기호로 표현되며 기하학적인 길이나 폭은 가지고 있지 않다. 점사상은 사상들의 특정한 위치를 가리키는데 대개 실제 사상들의 크기 그대로 표현하지 않고 축척에 따라 적절하게 나타난다. 또한 점은 주소나 교통사고 발생지점 등과 같은 무형의 사상들에 대해서도 위치를 나타낼 수 있다.

점사상에 필요한 정보는 위치좌표와 속성값이다. 수치지도에서 점사상은 자신의 위치에 부호로 표현된 사상이다. 점사상에는 주택, 건물과 같이 실재의 위치좌표를 갖는 사상이 주로 포함되어 있다. 일반 지도의 경우 논, 밭, 숲과 같이 일정한 면적을 갖는 지역을 점사상의 패턴(screen tone)으로 표현하고 있다. 이와같이 점 기호로 표현하는 경우는 점사상이라 할 수 없다.

점사상은 다른 점들과 함께 특정한 분포패턴을 나타내고 있다. 분포패턴에는 임의적(random) 분포, 규칙적(regular) 분포, 군집(clusterd) 분포 등의 세 가지가 있다. 이러한 점의 분포패턴을 분석하여 점사상의 분포특성을 파악하고 그 분포특성을 효과적으로 유지할 수 있도록 일반화 하는 것이 효과적이다. 본 연구에서는 건물을 대상으로 점사상의 분포패턴을 분석하고 그 분포패턴을 유지할 수 있도록 점사상 일반화 방안을 모색하였다.

3. 점사상 일반화 기법

3.1 일반화 기법

동일한 지역을 다른 축척의 수치지형도로 각각 제작한다면 지도정보의 중복과 제작비용의 비효율

적인 투자를 초래할 수 있다. 따라서 하나의 대축척 지도정보로부터 작은 축척의 지도정보를 추출할 수 있는 기법이 필요하다. 이러한 기법을 일반화라 한다. 지도가 제도사의 수작업에 의해 제작되던 과거에는 일반화기법도 제도사의 주관적 판단이 많이 개입되었다. 그러나 컴퓨터 기술의 발달과 함께 수치지형도가 등장하면서, 체계적이고 수학적인 일반화 알고리듬이 개발되기 시작하였다.

일반화는 축척의 변화에 따라 지도사상에 대하여 발생할 수 있는 단순화(simplification), 결합(merging), 변위(displacement), 통합(aggregation) 등을 포함한다. 이중에서 단순화는 일반화 과정에서 가장 핵심기술로 여겨지고 있다. 축척에 따라 지도정보의 내용을 시각적으로 다르게 표현할 뿐만 아니라 수치지형도의 정보양을 축척에 따라 조절할 수 있기 때문이다. 본 연구에서는 일반화의 의미를 단순화로 한정하고자 한다.

지도학에서 단순화(simplification)는 폭넓은 의미에서 '자료의 중요한 특성들을 결정하고 이를 가능한 한 보존하며 표시하고자 하는 축척에서 필요이상으로 상세한 지도사상이나 지도사상의 일부분(unnecessaries)을 제거함'을 가리킨다. 컴퓨터 지도학(computer-assisted cartography) 혹은 GIS에서 단순화의 필요성은 다음과 같은 맥락에서 부각된다. 즉, 대축척에서 취득된 지도 사상의 자료를 그보다 소축척으로 표현코자 할 때 기기(device)의 제한된 해상력으로 원자료의 상세도를 유지하지 못할 뿐만 아니라 불필요하게 긴 처리시간을 소요하게 된다. 또한 어떤 경우 지나치게 상세한 것이 오히려 중요한 정보에 대한 전달, 인식과 판단을 방해할 가능성이 있다.

지금까지 많은 연구들 통해 약 100여개의 선형사상의 단순화 알고리듬이 개발되어 있다. 이 알고리듬들은 가능한 한 원래의 기하학적 특성을 최대한 보존하여 원자료의 공간적 정확도 혹은 해상력(spatial accuracy or spatial resolution)을 최대로 유지하려는 측면과 가능한 처리속도를 빠르게 하려는 측면(computational efficiency)으로 정리된다.

이에 비하여 점사상에 대한 단순화 알고리듬은 거의 없는 실정이다. 기하학적인 모양을 가지고 있는 선형사상과 달리 점사상은 단순히 위치정보만을 가지고 있어 주요 연구 대상으로 다루어지지 않아 왔다. 그러나 점사상 역시 선형사상과 함께 지도정보를 구성하는 주요한 사상이다. 또한 점사상은 주의의 사상들과 함께 특정한 분포패턴을 지니고 있다. 점사상을 단순화 할 때 이러한 패턴특성을 무시한다면 지도정보의 왜곡을 초래할 수 있다. 따라서 점사상의 분포패턴을 분석하여 패턴이 유지되도록 일반화하여야 한다.

3.2 점사상의 일반화 방안

점사상의 일반화에 대한 연구방향은 크게 두가지로 나눌 수 있다.

첫째는 기준점을 중심으로 일정거리내에 다른 점사상이 위치하고 있으면 제거하는 방법이다. 이 방법은 알고리듬이 간단하고 명료하여 쉽게 구현할 수 있으며, 실질적으로 가장 널리 쓰이고 있는 방법이다. 그러나 단순히 점사상을 제거할 경우 점의 분포패턴이 변화할 수 있다.

둘째는 군집되어 형성된 점들을 하나의 면사상으로 표현하는 방법이다. 이 방법은 군집된 사상의 주위에 경계선을 설정하고, 경계선 내부의 점들을 하나의 등질적인 면으로 표현하는 것이다. 이방법은 축척의 변화가 클 경우, 즉 대축척에서 매우 작은 소축척으로 변화할 때 여러 사상들을 묶어 하나의 사상으로 표현하기 때문에 효과적이다. 그러나 여러 점사상들을 등질적인 하나의 사상으로 일반화하기 때문에, 축척의 변화가 작을 경우나 점들이 산재하여 분포할 경우에는 이 방법의 적용은 지도정보의 왜곡을 초래할 수 있다.

점사상에 대한 두가지의 일반화 방법 중에서 첫번째 방법이 가장 일반적이고 널리 이용되고 있는 방법이다. 본 연구에서는 첫 번째 방법, 즉 일정거

리내의 점들을 생략하는 기법을 점사상 일반화의 알고리듬으로 설정하였다. 일정거리내의 점들을 생략하는 기법은 점사상의 분포특성을 분석하고 이를 유지할 수 있도록 거리를 조절한다면 효과적인 점사상 일반화 기법으로 이용될 수 있다.

3.3 점사상 일반화의 고려사항

점사상을 일반화할 때는 양과 질의 측면에서 고려하여야 한다. 첫째로 양적인 측면에서 일반화의 효과가 나타나기 위해서는 적당한 양으로 점사상의 수가 줄어들어야 한다. 어느정도 정보의 양이 줄어들어야 사용자의 축척에 맞는 정보가 되기 때문이다. 축척에 따라 생략되는 적절한 정보양을 측정할 수 있는 축척 지표로 Töpfer의 법칙(Radical Law) 이용한다. Töpfer 법칙은 상대적으로 대축척인 자료원으로부터 원하는 소축척의 자료를 추출하기 위해 일반적으로 사용되는 법칙이다. 즉, 대축척 지도에서 나타나는 사상의 수와 상대적으로 소축척인 지도에서 나타나는 사상의 수에는 다음과 같은 관계가 있음을 나타내는 것이다.

$$N_f = N_a \times \sqrt{M_a / M_f}$$

Na : 대축척(Ma)에서 사상의 수
Nf : 소축척(Mf)에서 사상의 수

따라서 점사상을 일반화할때는 이러한 비율로 정보량이 감소하도록 고려하여야 한다.

둘째로 질적인 측면에서 점의 분포패턴이 유지되도록 일반화하여야 한다. 점사상은 선형사상과 같은 기하학적인 모양은 가지고 있지 않지만, 주위의 사상들과 어울려 특정한 분포패턴을 가지고 있다. 점사상을 생략할 때, 이러한 분포패턴을 무시한다면, 이는 선형사상에서 선의 모양을 무시하는 것과 같은 효과를 가지게 된다. 따라서 점의 분포패턴을 분석하고 이를 유지할 수 있도록 일반화가 이루어져야 한다.

4. 점사상의 분포특성과 분석기법

4.1 점사상의 분포패턴

일반적으로 점패턴의 기본 유형은 규칙적(regular), 임의적(random), 군집적(clustered) 유형으로 나누어진다. 이들은 각기 특정한 확률분포에 의해서 분포유형을 형성한다. 이들 분포유형은 관찰된 실제 점패턴을 비교·분석하는 기준이 되고 있으며, 또한 임의적인 분포유형으로부터 어느 정도 벗어나고 있는가를 설명해 주는 개념적 기틀이 되고 있다.

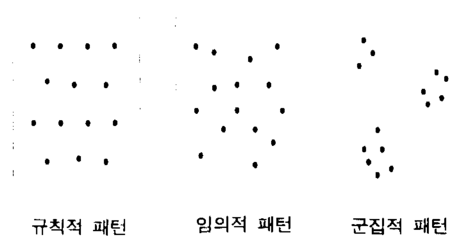

그림1 점사상의 분포패턴

그림에서 볼 수 있는 바와 같이 규칙적(regular) 패턴이란 점사이의 간격이 규칙적으로 배열된 분포를 나타내며, 군집적(clustered) 패턴은 점들이 밀집화되어 있는 배열상태에 따라서 각 점과 이웃하는 점과의 간격이 매우 좁게 나타난다. 임의적(random) 패턴이란 점사이의 간격이 임의적으로 나타난다. 따라서 일부의 점들은 조밀하게 분포하기도 하며 또 어떤 점들은 분산되어 있다.

흔히 군집화된 패턴과 규칙적 패턴은 시각적으로 쉽게 구분되지만 임의적 패턴은 시각적으로 판단하기 어려운 경우가 많다. 따라서 패턴 유형의 판단을 위해서 통계적인 분석방법이 요구된다.

4.2 점사상의 분포속성

점분포패턴은 두가지의 분포특성을 가지고 있다. 하나는 단위면적당 점의 수, 즉 점의 밀도이다.

점의 밀도가 지역에 따라 차이가 없다면, 그 점사상의 분포는 규칙적이라 할 수 있다. 반대로 지역에 따른 밀도의 차이가 크다면, 그 분포는 군집적이라 할 수 있다. 이와같이 점사상의 분포특성을 일정한 공간에서 점의 수로 나타낼 수 있으며 이를 점분포의 일차속성이라 한다. 점분포의 일차속성을 분석하기위한 가장 일반적인 분석방법이 방격분석(quadrat analysis)이다.

두 번째는 점간의 상호작용이다. 공간상에 표현된 점의 패턴이 점간의 상호작용에 의해 형성된 것인지, 임의로 분포한 것인가를 분석하는 것이다. 점들이 서로 연관되어 특정한 분포패턴을 형성하였다면 그 분포는 밀집, 혹은 규칙적이라 할 수 있다. 반대로, 혹은 점들이 서로 독립적으로 분포패턴을 형성하였다면 그 분포는 임의적이라 할 수 있다. 이와같이 특정분포가 점들간의 영향에 의하여 발생하였는가를 분석할 수 있으며 이를 점분포의 이차속성이라 한다. 이차속성은 점들간 공간적 상호작용에 대한 분석을 행하기 때문에 일차속성보다 점들간 공간적 상호의존성(spatial dependence)를 잘 파악할 수 있다. 점분포의 이차속성을 분석하기 위한 가장 일반적인 분석방법이 최근린분석(Nearest-Neighbour Analysis)과 K함수이다.

4.3 점분포 패턴분석기법

1) 방격분석(quadrat analysis)

점사상의 분포패턴을 분석하는 방법 중 가장 간단한 것으로, 연구대상지역을 등간격의 격자로 나누어 각 방안안에 놓여 있는 점의 수를 세는 분석방법이다. 정방형으로 분할된 각 방안안의 점을 세어서 도수분포표로 정리하면 공간상의 분포패턴을 어느정도 알 수 있다. 또한 도수분포표로 부터 점분포의 평균과 분산을 산출하여 어떠한 분포유형인가를 파악할 수 있다.

m의 격자에서 점의 수를 (xm)이라 할 때, 임의성을 표현하기 위한 통계적 검정법은 이러한 도수가 포아송 분포를 따른다는 점에서 출발한다. 즉 점의 수의 평균과 분산이 같으리라 가정한다. 이를 통계적으로 나타내면 다음과 같다.

$$X^2 = \frac{(m-1)s^2}{\bar{x}} = \frac{\sum_{i=1}^{m}(x_i - \bar{x})^2}{\bar{x}}$$

- m : 격자의 수
- $\bar{x}$: 격자의 평균 점의 수
- x_i : i번째 격자의 점의 수
- s2 : 분산

방격분석을 통하여 관측, 계산된 X2값과 χ^2 m-1 분포를 비교하여 유의도를 검정한다. 유의성이 크거나 작을 때 임의적 패턴으로부터 군집되었거나 혹은 규칙적인 패턴으로 판정할 수 있다.

$s^2/\bar{x}$를 분산정도(index of dispersion)라 한다. 임의적 패턴에서는 분산정도가 0의 값을 기지며, 군집된 패턴에서는 분산정도가 0보다 크게 나타난다. 또한 규칙적 분포에서는 분산정도가 0보다 작다.

분산정도를 이용하는 장점은 점패턴의 표본추출과 결합하여 이용될 수 있다는 것이다. 이 경우 m 격자는 대상지역에 임의로 산재하고, 각 격자마다 점의 수가 세어진다. 또한 이러한 표본추출 기법을 이용하여 밀도(λ)를 측정할 수 있다. 밀도값이 상수이고 임의의 패턴이라고 가정한다면, 격자의 면적이 Q일 때, 임의의 격자 수에대한 밀도는 다음과 같이 구할 수 있다.

$$\lambda = \frac{\bar{x}}{Q}$$

방격분석을 적용할 때 고려해야할 사항은 격자의 크기를 얼마로 선정하는가이다. 격자의 크기에 따라서 분포패턴 분석의 결과가 다르게 나올 수 있기 때문이다. 많은 연구에서 격자의 적당한 크기에 대하여 다루어 왔다. 생태학자의 경우

2A/N(A는 대상지역의 면적, N은 점의 수)가 적당하다고 여기고 있으며, 지리적인 연구에 있어서 공간적 경쟁이 심한 경우에는 A/N 까지도 적절한 크기로 보고 있다.

방격분석을 이용하여 여러 가지 분석이 가능하다. 그 중 대표적인 방법이 Grieg-Smith 방법이다. 이 방법은 원래의 격자크기에서 각각의 분산을 계산한 후 몇 개의 격자를 합쳐가며 하위 격자를 형성하여 다시 분산을 계산한다. 각 격자의 크기마다 분산을 측정하고 이를 격자크기에 대하여 그래프를 그린다. 그래프에서 정점이나 저점이 패턴 축척(scales of pattern), 즉 군집분포나 규칙적 분포의 징후로 해석한다. 이외에도 여러 가지 방격분석의 변형방법이 제시되고 있다.

방격분석의 문제점은 격자의 상대적 위치나 격자내에서 점의 상대적인 위치에 대한 고려가 없다는 점이다. 여기에서는 대상지역 전체의 모든 점들이 세어지며, 연속적인 격자로 이루어진 하나의 큰 격자로 간주하기 때문에 격자의 수와 각 격자의 상대적 위치에 대한 정보는 고려되지 않는다.

2) 최근린분석(Nearest-Neighbour Analysis)

최근린분석법은 어떤 현상의 공간적 분포를 결정하는데 지리공간상에서 가장 가까운 두 지점간의 거리를 측정하여 분포패턴을 기술한다는 점에서 방격분석법에 비해 보다 지리적인 접근방법이라고 볼수 있다.

최근린분석법은 어떤 임의의 점으로부터 다른 점들까지의 거리를 측정하고 그 결과에 근거하여 각종 평균거리를 산출한 후 실제로 관측된 점 분포패턴이 이론적인 점 분포패턴에 비해 어느 정도 벗어났는가를 파악하는 것이다. 최근린분석법의 기본적인 분석과정은 다음과 같다. 어떤 임의의 점으로부터 다른 점들까지의 거리를 측정한다. 각 지점에서 가장 가까운 다른 지점까지의 거리를 평균하여 관측된 평균 최근린거리를 산정한다. 점의 밀도를 이용하여 점 분포패턴으로부터 기대되는 평균 최근린거리($=1/2\sqrt{(n/A)}$: n은 점의 수, A는 대상지역의 면적)를 산정한다. 두 개의 평균 최근린거리를 비교하여 실제 관측된 점분포패턴이 이론적인 점분포패턴에 비해 어느 정도 벗어났는가를 파악한다.

기대되는 평균 최근린거리에 대한 관측된 평균 최근린거리의 비율을 최근린지수(nearest neighbour index : R)라 한다. 이 지수는 점분포패턴에서 각 점들의 간격이 임의적인 점분포패턴에서의 간격에 비해 어느정도 차이가 있는가를 나타낸다. R이 1인 경우는 완전히 임의적인 분포유형이고 R이 1보다 큰 경우에는 규칙적인 분포유형이고 R이 1보다 작은 경우에는 군집적인 분포유형이다. 따라서, R을 산정하고 그것이 1과 유의한 차이가 있는지를 판단하면 점의 분포가 어떤 유형의 분포인지를 결정할 수 있다.

R과 1사이의 유의한 차이를 검증하기 위해 Z검정법을 이용한다. 검정에 사용되는 표준오차(SEd)와 검정통계량(Z)은 다음공식을 이용하여 구할 수 있다. 여기서 SEd는 기대되는 평균 최근린거리의 표준오차이며 기대되는 최근린거리의 분산을 점의 개수 n으로 나눈 평방근이다.

$$\text{SEd} = \frac{0.261}{\sqrt{n^2/A}}$$

◦ n : 점의 수

◦ A : 대상지역의 면적

$$Z = \frac{\text{관측된최근린거리} - \text{기대되는최근린거리}}{\text{표준오차}(SE_d)}$$

점패턴 분포의 분석에서 점들이 공간적으로 상호작용하는가, 즉 점분포의 이차속성을 가지고 있는가를 측정하는 방법은 최근린 거리의 분포를 탐색하는 것이다. 최근린 거리는 작은 규모에서 점간 상호작용을 측정할 때 사용된다. 이를 표현하는 대표적인 방법이 G함수이다. G함수는 점간의

최근린 거리 W를 이용하여 경험적인 누적확률분포를 측정하는 것이다. G 함수는 다음과 같이 구할 수 있다.

$$\hat{G} = \frac{\neq(w_i \leq w)}{n}$$

◦n : 점의 수

◦#(wi ≤ w) : 기준거리 w보다 거리가 가까운 점의 갯수

G함수를 거리에 따라 그래프로 표현하면, 일정거리까지 급격한 상승곡선을 나타내며, 그 이후로는 평활한 모습을 보이게 된다. 상승곡선이 나타내는 지점까지는 최근린 거리가 짧은 경우가 많이 나타난다는 것을 뜻한다. 즉, 군집된 형태를 나타내고 있다. G함수는 최근린 분석법을 이용하여 거리에 따른 분포패턴 변화를 가늠할 수 있는 지표로 사용될 수 있다.

최근린분석법은 대부분이 주어진 점분포패턴이 어떤 분포유형인가를 기술하기 위한 목적으로 이용되고 있다. 또한 시간의 흐름에 따른 R 값의 변화추세로 분포패턴의 과정을 추론하기 위해 사용될 수도 있다. 그러나 연구지역의 범위에 따라 R 값이 변할 수 있으므로 범위가 다른 두 지역의 분포패턴을 비교할때는 주의하여야 한다.

3) K함수

점분포의 이차 속성을 분석하기 위한 또다른 방법은 K함수이다. K 함수는 점자료의 공간적 배열상태를 평가하기 위한 기법으로서, 점자료의 분포상태가 공간적 상호작용에 의해 형성되었는지 여부를 설명하는 통계적 기법이다

K함수는 특정지점으로부터 일정거리내의 실제적으로 분포하는 점의 수와, 이론적으로 규칙적인 점의 수를 비교함으로서, 점의 분포가 임의적인가 판정하는 함수이다. 구체적인 식은 다음과 같다.

$$\hat{K}(h) = \frac{1}{\lambda^2 R} \sum\sum I_h(d_{ij})$$

◦R : R 지역의 면적

◦ λR : R지역의 평균 event 발생 수

◦ d_{ij} : R 지역내에 위치한 i번째 event와 j번째 event간의 거리

◦ $I_h(d_{ij})$: indicator function (d_{ij} < h이면 1, 그렇지 않으면 0)

임의의 패턴일 경우 각 지점에서 점이 있을 확률은 동일하며 서로 독립적이다. 따라서 표본추출된 한 지점으로부터 일정한 거리 h이내에서 발견될 평균 점의 수는 $\lambda\pi$h2 이다. 즉, 공간적 상호작용이 존재하지 않는 등질적 상황에서 K(h) = πh2 라고 할 수 있다. 한편, 군집된 분포일 경우 K(h) > πh2이고, 규칙적 분포일 경우는 K(h) < πh2 로 나타난다.

이러한 이론적 근거를 바탕으로 πh2와 표본자료에서 구한 K(h)를 비교함으로써 점자료의 패턴을 탐색할 수 있다. K 함수를 비교하기 위하여 πh2의 역으로 환산한 함수가 L함수이다. L 함수는 다음과 같이 구해진다.

$$\hat{L}(h) = \sqrt{\frac{\hat{K}(h)}{\pi}} - h$$

L함수의 값이 0보다 클 경우 군집된 패턴이며, 0보다 작을 경우 규칙적 형태를 나타낸다. L 함수를 거리에 따라 그래프로 표현할 경우, 원점에서부터 기울기가 1인 직선이 임의의 패턴을 나타낸다. 실제의 점의 분포가 군집된 형태일 경우 L 함수는 직선의 상단에 위치하며, 규칙적 형태일 경우 직선의 하단에 위치하게 된다.

5. 사례지역 분석과 실험

5.1 사례지역

본 연구에서 사용한 자료는 공주지역의 1 : 25,000 지형도이다. 1 : 25,000 지형도에서 건물 사상만을 선정하여 디지타이징 한 후, 2Km×2Km 크기의 사례지역을 추출하여 분석에 이용하였다. 그림 2는 본 연구에 사용된 사례지역의 점분포이다. 분석에 이용된 점사상은 모두 512개이며, 대상지역의 면적은 4,000,000m2이다.

본 연구에서는 실험적으로 1 : 25,000 축척의 건물사상으로부터 1 : 50,000 축척의 건물사상을 추출하는 일반화 알고리듬을 적용하고자 한다.

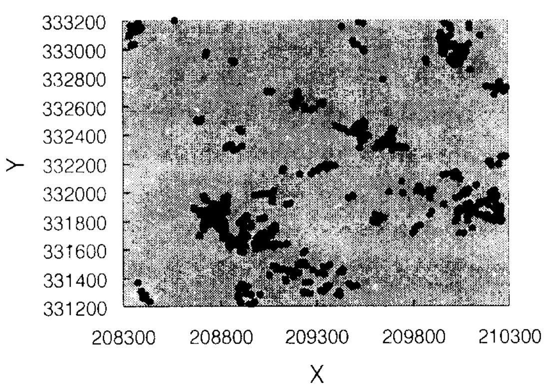

그림2 사례지역의 건물사상 분포

5.2 원자료의 분석

1) 방격분석

사례지역의 일차속성, 즉 점의 밀도를 분석하기 위하여 방격분석을 적용하였다. 먼저 격자의 크기에 따른 분포패턴을 보기 위하여 Grieg-Smith 방법을 원용하였다. 5m 격자로부터 5m씩 격자의 크기를 늘려가면서 분산정도 값을 비교하였다. 분산정도는 분산과 평균의 비율로서, 대상지역의 분포형태를 나타내는 척도로 이용된다. 분산정도의 값이 크면, 격자간 돗수의 차이가 크다는 것을 의미하며, 따라서 군집된 형태로 해석할 수 있다. 반대로 분산정도의 값이 작으면 규칙적 분포로 해석할 수 있다. 사례지역을 대상으로 격자의 크기에 따라 방격분석을 한 결과 그림과 같은 분산정도의 변화를 볼 수 있다. 격자의 크기가 증가할수록 분산정도 역시 증가하고 있다. 격자가 커질수록 군집의 성격을 더욱 강하게 보이고 있다. 이는 사례지역의 자료가 군집된 형태를 가지고 있다는 것을 의미한다.

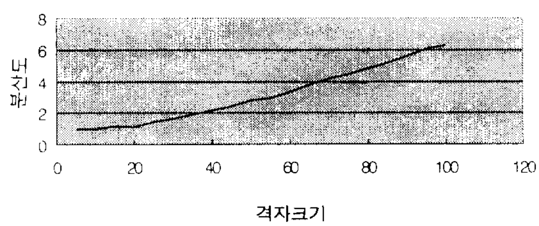

그림3 격자크기에 따른 분산정도의 변화

분산정도의 값은 포아송 분포의 분산정도와 비교하여 그 유의성을 검정한다. 포아송 분포는 평균과 분산의 값이 모두 1인 특징을 가지고 있다. 따라서 분산정도의 값이 1이면 임의적인 분포로 볼 수 있다. 분산정도의 값이 1보다 작아질수록 규칙적 분포유형에 가까우며, 1보다 커질수록 군집된 분포유형에 가깝다고 할 수 있다. 임의적 분포유형을 뜻하는 분산정도 1로부터의 편차가 어느 정도 유의한가를 조사하기 위하여 그 차이에 대한 표준오차를 구한 후 검정법을 사용한다. 표준오차를 구하는 식은 다음과 같다.

$$표준오차 = \sqrt{2/(N-1)} \quad (N : 점의\ 수)$$

검정에 필요한 기대값은 다음과 같이 구할 수 있다.

$$Z = \frac{관찰비율 - 기대비율(1)}{표준오차}$$

Z 기대값을 계산한 후, 유의수준에 적당한 임계

치와 비교하여 유의성을 검정한다. 그림 4는 격자크기에 따른 분산정도의 기대값을 그래프로 표현한 것이다. 사례지역의 경우 15m 격자크기부터 Z값이 증가하고 있으며, 그 값도 99% 유의수준의 임계치인 2.58보다 크다. 따라서 군집된 패턴을 나타내고 있다. 그림에서 20m 격자의 크기에서 기대값이 감소하는 현상을 볼 수 있다. 이는 20m 크기의 격자에서 어느정도 패턴의 변화가 있다는 것을 나타낸다.

사례지역에 대한 방격분석의 결과 건물사상은 군집된 형태를 보이고 있으며, 격자의 크기가 커질수록 더욱 군집된 형태를 보이고 있다. 그러나 20m 격자크기에서는 약간의 패턴변화가 나타나고 있다.

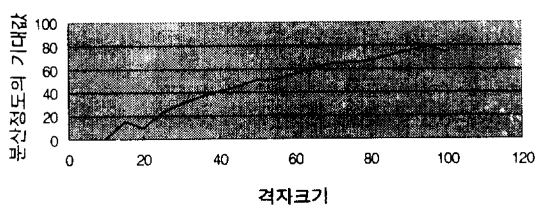

그림4 분산정도의 기대값 변화

2) 최근린분석

점분포의 이차속성을 파악하기 위하여 최근린분석을 적용하였다. 사례지역을 대상으로 최근린분석을 적용한 결과 최근린 거리는 20.166이며, 최근린지수(R)은 0.456으로 나타났다. R이 1에 가까울수록 규칙적 패턴이며, 1보다 작을수록 군집형 패턴을 나타낸다. 사례지역에서 측정된 최근린 지수와 임의적 패턴에서의 최근린 지수(1)와의 차이를 검정하기 위하여 계산되는 검정통계량은 23.567로서, 99% 유의수준에서 임계치(2.58)보다 크므로 유의싱이 있다. 사례지역의 건물사상은 최근린분석 결과 군집된 형태로 판명되었다.

거리에 따른 점간의 상호작용 정도를 측정하기 위해 G함수를 구성하였다. G 함수는 확률분포상에서 점간의 최근린 거리를 표현한 것이다. 그림 5는 사례지역의 G함수를 표현한 것이다. 그림에서 30미터 거리까지는 급격한 증가를 보이다가 그 이후에는 평활한 곡선이 나타나고 있다. 이는 30미터 거리까지 군집된 형태를 보이고 있다가 그 이후에는 분포패턴이 변화하게 되는 것을 나타낸다.

사례지역에 대한 최근린분석을 한 결과 건물사상은 군집된 형태로 나타났다. 또한 사례지역의 G함수를 구한 결과 30m거리까지는 군집된 패턴을 보이고 있으며, 30m이후에는 패턴의 변화를 보이고 있다. 이는 30m거리까지는 군집된 패턴이 유지되고 있음을 나타낸 것이다.

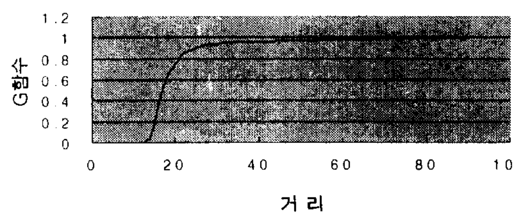

그림5 최근린 거리(G 함수)

3) K함수

K함수는 사상간의 이차속성을 분석하는 방법이다. 사례지역을 K함수로 표현한 후, 이를 표준화한 L함수로 나타내면 그림 6과 같다. 그림에서 대각선 직선이 임의의 패턴에서의 L함수를 나타낸 것이며, 곡선은 사례지역에서 측정된 L함수를 나타낸 것이다. 그림에서 거리가 낮은 지점은 곡선이 직선의 상단에 위치해 있다. 이는 작은 거리에서는 군집된

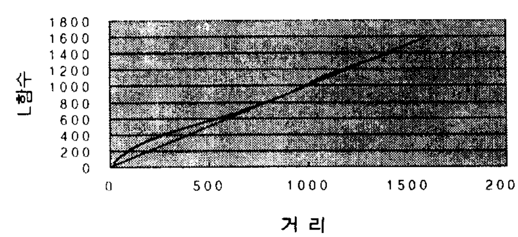

그림6 L 함수

패턴을 보이고 있으며 축척이 커질수록 임의적 분포로 패턴이 변화하는 것을 알 수 있다.

4) 원자료의 분석결과

사례지역의 건물자료를 이용하여 방격분석과 최근린 분석을 적용하여 점의 분포패턴을 분석하였다. 그 결과 모든 패턴분석에서 연구대상지역은 군집된 패턴으로 판명되었다. 또한 분석의 규모거리가 증가함에 따라 점차로 패턴이 군집된 패턴에서 임의적 패턴으로 변화하는 것을 알 수 있다. 방격분석과 최근린 분석의 결과 20m에서 30m 가이의 거리에서 패턴의 변화가 나타나고 있다. 따라서 사례지역의 점사상을 일반화할 때 단순화되는 일반화 거리를 20m에서 30m 이상으로 설정할 경우 점분포패턴의 변화가 올 수 있다는 것을 암시한다. 다음 으로는 실제로 사례지역의 자료를 대상으로 단순화 거리를 조절하면서 패턴이 변화하는 양상을 살펴보고자 한다.

5.3 일반화된 점사상의 분석

1) 사례지역의 점사상 일반화

본 연구에서는 점사상의 일반화 방법으로 점의 밀도를 조정하는 방안을 택하였다. 점간의 거리가 설정된 임계치보다 작으면 해당되는 점을 생략한다. 본 연구에서는 점사상을 생략하기 위해 설정된 임계치를 일반화 거리로 칭하기로 한다. 점간의 일반화 거리가 크면 많은 양의 점이 생략되며, 대축척에서 소축척으로의 큰 규모의 축척변화가 나타난다. 그러나 너무 큰 일반화 거리를 설정하게 되면, 점분포패턴이 변화하게 되어, 점사상이 원래 가지고 있던 분포특성이 상실된다. 따라서 적당한 크기의 일반화 거리를 설정하여야 한다.

본 연구에서는 사례지역의 건물사상을 대상으로 5m에서 50m까지 5m씩 일반화 거리를 늘려가며 일반화하였다. 각 거리에 따라 일반화된 점사상을 대상으로 일반화된 점사상의 수량을 분석하고 패턴분석을 적용하여 일반화의 양에 따른 패턴의 변화를 탐색하였다.

2) 일반화된 점사상의 수량 분석

그림 7은 각각의 일반화 거리에 따라 일반화된 점사상의 수를 나타내고 있다. 그림에서 20m 거리까지는 급격한 하향곡선을 그리다가 그 이후에는 비교적 완만한 감소를 보이고 있다. 한편, 축척에 따른 점의 수를 가늠할 수 있는 Radical Law를 적용할 경우 1 : 25,000의 축척으로 부터 1 : 50,000의 축척에 해당되는 점의 수는 다음과 같다.

$$512 \times \sqrt{50,000/25,000} = 512 \times \sqrt{2} = 362$$

Radical Law에 의하여 계산된 점의 수인 362tro에 맞춘 사례지역의 일반화 거리는 16m에 해당된다. 점의 수를 기준으로 하였을 때, 16m의 일반화 거리를 주는 것이 바람직하다.

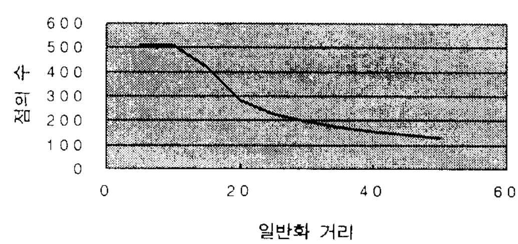

그림7 일반화 거리에 따른 점의 수 변화

3) 일반화된 점사상의 패턴분석

가. 방격분석

일반화 거리에 따라 각각 일반화된 점사상 자료를 대상으로 50m×50m의 격자를 씌운 후, 방격분석을 적용하였다. 그림 8은 방격분석을 적용하여 측정된 분산정도 즉, 분산과 평균의 비율을 나타낸 것이다. 그림에서 일반화 거리가 증가할수록 분산정도는 점점 감소되어 군집된 패턴이 사라지

고 임의적 패턴에 다가가고 있다. 또한 20m 지점을 기점으로 곡선의 기울기가 완만하게 변화하고 있다. 이는 일반화되는 거리가 늘어나면서 군집적 패턴에서 임의적 패턴으로 분포양상이 바뀌고 있으며, 특히 20m이상의 거리부터 패턴의 변화가 나타나기 시작한다는 것을 의미한다.

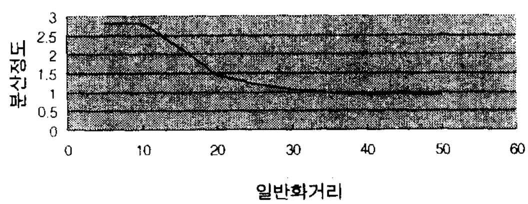

그림8 일반화 거리에 따른 방격분석의 분산정도 변화

유의성 검정을 위하여 분산정도의 값을 기대치로 변환하여 관측한 결과는 그림 9와 같다. 기대값의 그래프 역시 일반화 거리가 증가할수록 기대값이 낮아지고 있으며, 일반화 거리가 30m 이후부터는 99%의 유의성 검정량인 2.58보다 작은 값을 보이고 있다. 일반화 거리가 30m 이상되는 점사상의 일반화 결과는 통계적으로 군집된 패턴이 아니라고 판정할 수 있다. 또한 기대값의 그래프에서도 20m 거리에서부터 기울기가 변화하고 있다. 이는 20m 거리를 패턴이 변화하는 기준점으로 해석할 수 있다.

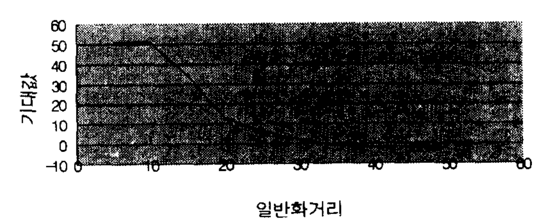

그림9 일반화 거리에 따른 분산정도의 기대값 변화

나. 최근린 분석

사례지역의 건물사상을 대상으로 5m부터 50m까지 5m 간격으로 일반화 거리를 설정하여 일반화한 자료를 최근린 분석한 결과는 그림 10과 같다. 그래프는 일반화 거리에 따라 일반화된 점사상 자료의 최근린 지수를 측정하여 표현한 것이다. 최근린 지수는 실제의 최근린 거리와 이론적인 최근린 거리와의 비율로서 지수의 값이 1에 가까울수록 임의적 분포패턴을 나타낸다. 그래프와 같이 일반화 거리가 증가할수록 최근린 지수는 1의 값으로 접근하고 있다. 또한 20m를 기점으로 기울기가 변화하고 있다. 이는 일반화 거리가 증가할수록 임의의 분포패턴으로 패턴이 변화하고 있으며, 20m 거리부터 거리에 대한 변화의 속도가 달라지고 있음을 알 수 있다. 즉, 20m의 일반화 거리를 기준으로 분포패턴이 변화하고 있음을 나타내고 있다.

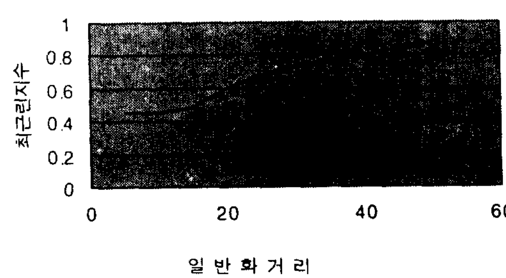

그림10 일반화거리에 따른 최근린지수의 변화

그림 11은 최근린 지수의 유의성 검정을 위해 표준오차를 개입한 기대값의 그래프이다. 기대값의 그래프에서도 일반화 거리가 증가할수록 기대값이 점점 작아지고 있으며, 50m의 거리부터는 99% 유의수준의 임계치 2.58 보다 작은 값을 가지고 있다. 일반화거리가 50m이후에는 통계적으로 임의의 분포패턴이 된다고 할 수 있다. 그래프에서 20m를 기점으로 곡선의 기울기가 완만하게 변하고 있다. 이는 일반화 거리가 20m부터 분포패턴이 변화하고 있음을 나타내고 있다.

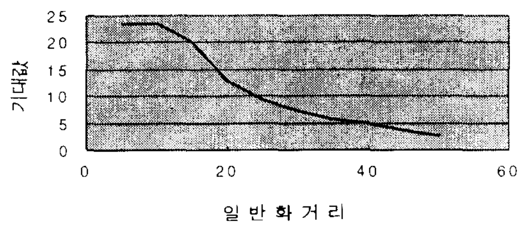

그림11 최근린지수의 기대값 변화

4) 일반화된 점사상의 분석 결과

사례지역의 건물사상을 대상으로 일반화 거리를 설정하여 분포패턴을 분석한 결과 다음과 같은 특성을 찾을 수 있었다. 첫째로 일반화 거리가 증가할수록 분포패턴은 군집된 형태에서 임의적 형태로 변화한다. 일반화 거리가 증가하게 되면, 많은 양의 점들이 생략되며, 따라서 점사상의 분포형태도 변화하게 된다. 이는 일반화대상이 되는 점사상이 군집된 패턴일 경우 특히 잘 나타난다. 둘째, 일반화 거리에 대한 패턴의 변화는 특정거리를 기준으로 그 변화양상이 달라지게 된다. 사례지역의 경우 20m를 기준으로 방격분석과 최근린분석에서 측정된 각 지수와 기대값 곡선의 기울기가 변화하였다. 이는 특정거리를 기준으로 분포패턴이 변화양상이 달라진다는 것을 의미한다.

사례지역을 대상으로 일반화를 적용한 결과 20m 거리가 점사상의 분포패턴이 변화하는 분기점으로 나타났다. 즉, 사례지역애서 일반화 거리를 20m 이상으로 설정할 경우, 원자료의 분포특성을 상실할 수 있다.

6. 결론과 제안

대축척 지도에서 소축척지도로 일반화하는 과정에서 점사상은 선형사상에 비하여 연구대상으로서 부각되지 못하였다. 그러나 점사상도 선형사상과 마찬가지로 특정한 분포패턴을 가지고 있다. 따라서 점사상을 일반화 할 경우에 연구자는 이러한 분포패턴이 변화하지 않도록 신중히 고려하여야 한다.

본 연구에서는 공주지역 일부를 사례지역으로 건물사상 대상으로 점사상 일반화 실험연구를 수행하였다. 실험연구의 결과는 다음과 같다.

첫째, 건물사상의 분포는 군집형태를 가지고 있다. 둘째, 건물사상은 관측거리에 따라 군집된 형태에서 임의적 분포로 패턴이 변화하고 있다. 셋째, 건물사상을 일반화할 때, 일반화 거리에 따라 분포패턴이 군집된 패턴에서 임의적 패턴으로 변하며, 20m 거리가 분포변화의 기준거리로 나타났다. 건물사상은 일반화 정도에 따라 분포패턴이 점차 변화하며, 특히 일정거리(20m)를 넘어서면 다른 유형의 분포패턴으로 변화하게 된다. 건물사상을 일반화 할 때에는 분포패턴이 변화하지 않도록 기준거리를 설정하여 그 이상의 일반화는 조심하는 것이 바람직하다.

본 연구의 결과를 토대로 향후 점사상의 일반화 과정에서 고려하여야 할 사항들을 제안하면 다음과 같다. 첫째, 점사상을 일반화하기에 앞서 그 분포특성을 분석하여야 한다. 점사상의 분포특성과 거리에 따른 특성변화정도에 따라 일반화 기법이 다르게 적용되어야 한다. 둘째, 점사상을 일반화 할 때에는 점사상의 양적인 측면과 질적인 측면 모두를 고려하여야 한다. 양적인 측면에서, 일반화 정도에 따라 자료량이 감소할 수 있도록 일반화 거리를 설정하여야 한다. 반면 질적인 측면에서, 원래의 분포패턴이 변화하지 않는 범위내에서 일반화 거리를 설정하여야 한다. 점사상의 양과 질이란 두가지 관점에서 적절한 일반화 거리를 모색하는 과정이 필요하다.

점사상은 선형사상과 함께 수치지형도를 형성하는 주요한 구성요소이다. 점사상의 일반화는 수치지형도의 일반화에서 큰 비중을 차지하고 있다. 점사상도 선형사상과 마찬가지로 일반화 기술에서

주요한 연구대상이다. 점사상을 일반화하면서도 그 분포패턴을 유지할 수 있는 다양한 연구가 계속되어야 할 것이다.

참 고 문 헌

1. 남영우, 1992, 계량지리학, 법문사.
2. Arninghaus, S. Lach, 1993, Pratical Handbook of Digital Mapping, CRC Press.
3. Bailey, Travor C. and Anthony C. Gatrell, 1995, Interactive Spatial Data Analysis, Longman Scientific & Technical.
4. Buttenfield, Batbara P. and Robert B. McMaster, 1991, Map Generalization, Longman Scientific & Technical.
5. Clark, P. J. and F .C. Evans, 1954, "Distance to nearest neighbor as a measure of spatial relationships in populations", Ecology, v 35, pp. 445-453.
6. Mathdoft, 1996, S+ Spatialstats User's Manual.
7. McMaster, Robert B. and K. Stuart Shea, 1992, Generalization in Digital Cartography, Association of American Geographers.
8. Muller, Jean-Ckaude, Jean-Philippe Lagrange and Robert Weibel, 1995, GIS and Generalization : Methodology and Practice, GISDATA 1, Taylor & Francis.
9. Robinson, A. H., R. D., Sale, J. L., Morrison, and P. C. Muehrcke, 1984, Elements of Cartography, (5th ed.), John Wiley & Sons.