

형태적 특성을 고려한 지형 모델

최미나^o 김시완 김학철 이기준

부산대학교 전자계산학과

{ mnchoi, swkim, hckim }@quantos.cs.pusan.ac.kr, lik@pusan.ac.kr

Morphological Surface Model

Mina Choi^o Si-Wan Kim Hak-Cheol Kim Ki-Joune Li

Dept. of Computer Science, Pusan National University

요 약

지리정보시스템이 널리 사용됨에 따라 여러 다양한 기능들이 요구되고 있다. 특히, 2차원이 아닌 지표면의 고도를 가진 지형데이터를 표현하고 분석하는 기능이 지리정보시스템의 중요한 기능으로 요구되어지고 있다. 현재 지형데이터를 표현하는데 가장 많이 이용되는 모델은 DEM, TIN, 등고데이터가 있는데 각각의 모델은 응용분야에 따라 장단점을 가진다. 지리정보시스템에서 사용되는 실제의 지형데이터가 대용량이기 때문에, 데이터를 분석하는데 드는 비용이 매우 클 뿐만 아니라 분석을 위한 데이터 로드에도 상당한 비용이 든다. 그러므로 최대한 적은 데이터 양으로 지형을 표현하는 것이 지리정보시스템에서는 매우 중요하다. 본 논문에서는 데이터 양을 줄임과 동시에 지형의 중요 정보를 최대한 보존하는 형태적 지형 모델을 제안한다. 실험에 따르면, 제안한 모델은 TIN에 비해 데이터의 양이 작을 뿐만 아니라 지형의 주요 형태적 정보를 유지함으로써 연산의 정확성도 높이는 장점을 보인다.

1. 서론

지리정보시스템이 널리 사용됨에 따라 여러 다양한 기능들이 요구되고 있다. 특히, 2차원이 아닌 지표면의 고도를 가진 지형데이터를 표현하고 분석하는 기능이 지리정보시스템의 중요한 기능으로 요구되어지고 있다. 현재 이러한 기능을 완벽히 지원하는 지리정보시스템은 거의 없는 실정이며 이러한 지형 데이터를 표현하고 이용하기 위한 연구가 점차 증가하고 있다.

지형 데이터를 분석하는 기능의 어려움은 다음의 2가지 사실에 기인한다. 첫째, 지형데이터는 대용량이므로 그 처리 비용이 높고 처리가 매우 난해하다. 실제로 NASA는 SRTM(Shuttle Radar Topography Mission)에서 지구 육지의 약 80%를 30미터 간격의 격자로 표현한 지형데이터를 만들었다. 이 데이터의 크기는 약 10테라 바이트이다. 현재 이러한 대용량 지형데이터를 사용할 수 있는 응용프로그램은 거의 없다. 따라서 대용량 데이터를 처리하는 것은 지리정보시스템 분야에서 매우 활발히 연구되고 있고 시간, 공간 데이터베이스 분야에서도 점차 관심이 커지고 있다. 둘째, 지형데이터에 대한 분석 기능은 단지 2차원 평면에서 위치 정보만 이용하는 일반적인 지리정보 시스템의 분석 기능에 비해서 알고리즘의 복잡도가 높다.

지형데이터는 동일한 지형에 대해서 표현 방법에 따라 저장 용량의 차이가 있다. 지형데이터를 표현하는 일반적인 방법들은 DEM(Digital Elevation Model), TIN(Triangulated Irregular Network), 등고 데이터 표현 방법 등이 있는데 각각의 표현 방법들은 특정한 응용 분야에 따라 장단점을 가진다. 이 중에서 현재 가장 널리 사용되는 것은 DEM 표현 방법인데 그 이유는 표현이 간단하고 분석이 쉽기 때문이다. 그러나 DEM으로 지형을 표현하는 경우에는 데이터의 양이 너무 많아지는 단점이 있다. 지형 데이터 표현 방법 가운데 가장 데이터 양이 적은 것은 TIN이다. TIN은 데이터의 양은 작지만, 분석 알고리즘이 매우 복잡한 단점이 있다. 그러나 TIN

으로 구성되는 경우에도 상대적으로 다른 방법에 비해서 저장 용량이 작지만 그 크기도 결코 적지 않다.

이를 고려하여 본 논문에서는 TIN에 비해 더 적은 양의 데이터를 가지며 원래 지형이 나타내고 있는 중요한 형태적인 특성을 최대한 보존하여 TIN에 비해 특정 분석 기능을 쉽게 구현할 수 있는 새로운 데이터 모델을 제시한다.

실험 결과에 따르면 본 논문에서 제안한 모델은 동일한 지형에 대해서 비슷한 저장용량을 가지는 TIN 모델에 비해서 분석이 용이하며 성능의 향상을 보여 주었다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 관련 연구를 3장에서는 제안한 모델에 대해 설명하고 4장에서는 제안한 모델의 타당성을 실험한다. 마지막으로 5장에서는 결론 및 향후 연구에 대해서 제시한다.

2. 관련 연구

지형데이터는 기존의 공간데이터와는 다른 특성을 가진다. 일반적인 공간데이터가 2차원의 값인데 반해 지형데이터는 고도 값이라는 추가적인 속성이 존재하는 3차원 데이터이다. 또한, 실제 지형이 연속적이므로 그 크기가 매우 커진다. 이러한 지형데이터의 특성 때문에 지형데이터의 표현방법[1,2]이나 색인 방법, 질의 처리 방법[4,5,6] 등이 기존의 공간데이터의 표현 및 처리 방법과는 달라야 한다. 그러므로 이것들에 관한 연구가 활발히 진행되고 있다. 또한, 대용량의 지형데이터 처리 효율을 높이기 위해 병렬 디스크 및 환경에 관한 연구 등이 이루어지고 있다[7].

지형데이터 표현을 위해 널리 사용되는 모델에는 TIN, DEM, 등고 데이터가 있다. 일반적으로 가장 널리 이용되는 모델은 DEM이다. 아래의 [표1]은 서술한 모델 가운데 실제 분석 시 이용되고 있는 DEM과 TIN의 특성을 요약한 것이다. [표1]에서 알 수 있듯이 DEM은 TIN에 비해 데이터 양이 많을 뿐 아니라 정확성이 떨어진다.

1리므로 연산의 빠른 처리와 정확성을 높이기 위해 새로운 제안으로 부각되고 있는 모델이 TIN이다.

[표1] TIN과 DEM의 특성

	TIN	DEM
데이터 표현	삼각형의 네트워크	고정 격자
데이터 양	타 모델에 비해 적다.	많다.
데이터 조작	복잡	단순
정확도	타 모델에 비해 정확	떨어짐

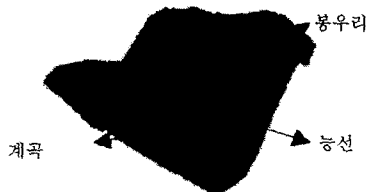
그러나 DEM에 비해 적은 데이터 양을 가진 TIN 역시 분석을 위해서 모든 데이터를 주기억장치에 적재해야 한다. 이 경우 지형 데이터 자체가 상당히 대용량이므로 주기억장치에 상주해야 하는 실제 TIN 데이터 양은 상당하다. 따라서 데이터를 디스크에서 읽어오는 시간이 길어질 뿐만 아니라 데이터의 크기가 크면 주기억장치에 모두 적재할 수 없다. 또 TIN 생성시 가장 널리 이용되는 들로니 삼각 분할 계열[1,2]로 구성된 TIN에서는 지형의 형태적인 중요 특성을 보존하기 어렵다. 왜냐하면, 단순히 점들 간의 기하관계를 이용하여 TIN을 구축하기 때문에 계곡이나 능선과 같은 실제 지형의 중요한 특성을 보존하지 못한다. 예를 들어 배수 분석에서 사용되는 배수 네트워크나 능선 네트워크와 같은 형태적인 특징이 제대로 나타나지 않는 경우도 있다. 다시 말해 TIN을 구성할 때 형태적 특성을 고려하지 않고 있기 때문에 중요한 속성을 놓치게 되어 분석의 장애가 되는 경우가 발생한다.

실제 응용 프로그램에서 지형 데이터를 분석하는 분야는 크게 경사도 및 적도량 분석, 가시성 분석[4], 경로 분석[5], 배수 분석[6], 지형 상세화 등 다양하다. 이러한 분석을 위해서 적절한 알고리즘과 자료구조가 필요하다. 이를 위해 계산 기하학(Computational Geometry)[3] 및 알고리즘 분야에서의 관련 연구가 이루어지고 있다[2].

3. 형태적 특성을 이용한 지형 모델

3.1. 지형에서의 형태적 특징

지형에서의 형태적인 특징이란 그 지형의 형태적인 특징을 설명하는 특정 속성을 말한다. 예를 들어 웅덩이 지형, 봉우리 지형, 능선, 계곡 등과 같은 속성을 형태적인 특징이라 할 수 있다[2]. [그림1]에서 지형의 형태적인 특징을 간략히 살펴볼 수 있다. 지형은 능선, 계곡 등과 같이 선이 지나는 속성이 있고 봉우리 안장점 등과 같이 점이 지나는 속성이 있다.



[그림 1] 지형에서의 형태적 특징

이와 같이 지형에 나타나는 특징이 분석을 위한 지형데이터로 만들 때 중요한 이유는, 지형데이터를 이용하여 사용자가 원하는 분석을 할 때 대부분의 분석 알고리즘이 위의 지형적 특성을 이용하기 때문이다.

예를 들면, 강수 유역권을 분석하거나 강물의 오염원을 찾는 경우 지형의 계곡과 능선 그리고 봉우리 등을 이용하여 분석하거나 찾는 방법이 있다[6]. 또한, 가시권 분석에[4] 능선을 이용하여, 경로 분석 알고리즘[5]은 안장점을 이용하기도 한다.

3.2. 형태적 특성을 고려한 지형 모델

본 연구에서는 원 지형이 가지는 계곡, 능선, 봉우리 등의 형태적인 특성을 최대한 보존하면서 기존의 모델에 비해 데이터 크기가 적은 새로운 방법을 제안한다.

형태적 특성을 고려한 지형 모델은 다각형의 집합으로 이루어져 있는데, 이 다각형은 원래 지형의 능선이나 계곡 정보와 같은 형태적 특징을 보존한다. 즉, 계곡과 같은 경우 하나 이상의 선으로 표현되어지는데 계곡을 나타내는 이러한 선들이 모두 하나 이상의 다각형에 포함되는 것이다. 이러한 방식으로 원 지형에서 선이나 점으로 나타나는 여러 특징을 지형데이터에 포함하는 것이다. 이러한 특징은 다시 응용프로그램에서 분석할 때 이용할 수 있다. 또한, 저장하는 단위가 다각형이므로 저장하는 단위가 삼각형인 TIN보다 같은 지형을 데이터로 만들었을 때 더 작은 크기가 된다.

본 연구에서 제안한 모델의 장점은 크게 다음의 두 가지로 요약된다.

- 적은 양의 데이터
- 지형의 형태적인 특성 보존

아래의 [그림2]는 실제 지형의 능선과 계곡의 두 가지 중요한 형태적 정보를 유지한 것으로 형태적 특성을 고려한 모델로 표현한 것이다. 즉, 앞서 언급한 바와 같이 원 지형의 계곡과 능선을 다각형을 이루는 변에 모두 포함시켰다. 이렇게 표현한 경우 TIN에 비해 적은 데이터 양을 가진다.



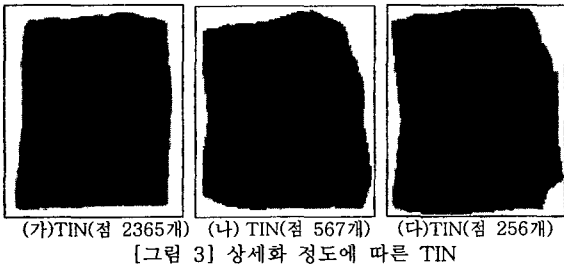
[그림 2] 형태적 특성을 고려한 지형 모델

본 연구에서 제시하는 형태적 특성을 고려한 지형 모델은 현재 수행할 수 있는 몇몇 분석에 한계가 있다는 단점이 있다. 예를 들어, 특정 위치의 고도 값을 추정한다거나, 경사도 및 적도량 분석, 계곡선과 능선을 추출해 내는 것과 같이 형태적 특성을 추출하는 분석은 간단할 뿐 아니라, 분석의 비용이 적은 장점을 가진다. 반면, 최적 경로 분석이나 배수 분석과 같은 복잡한 알고리즘을 요하는 분석에서는 다각형 단위로 분석하는 것과 관련된 알고리즘이 없기 때문에 까다로워진다. 이것은 TIN이 각각의 삼각형, 선, 점 등이 서로 위상관계를 가지는 반면 이 모델은 각각의 다각형 사이의 위상관계가 존재하지 않기 때문이다.

그러나 실제 모든 목적에 부합하는 지형 데이터 모델이 존재하기는 힘든 실정이다. 현재 일부 제한적인 기능을 지니지만 본 모델의 활용 분야도 많이 있다. 예를 들어, 선배이 항해를 할 때 특정한 지역에 암초가 있는가를 판단하는 매우 중요하다. 이 때는 특정한 지역의 고도 값 중 특정 값 보다 큰 것이 있는가를 분석하면 된다. 해저 지형은 바다가 넓은 만큼 매우 거대한 데이터가 될 것이므로 TIN이나 DEM을 사용하는 것 보다 이 방법을 사용한다면 해저 지형을 저장하기 위한 공간을 줄일 뿐 아니라 분석 시간을 많이 줄일 수 있을 것이다. 뿐만 아니라, 지형의 형태적인 특성을 추출하는 경우 상당히 간단하고 빠른 장점을 지닌다.

4. 실험

실험에서 사용한 데이터는 한국의 설악산 지형 데이터이다. 아래의 [그림3]은 설악산 지형 데이터를 TIN으로 구축한 데이터를 화면에 출력한 것으로 세 개의 TIN 데이터 모두 전체적인 지형의 모양이 유사함을 알 수 있다. 우측에 존재하는 TIN 데이터는 보다 자세한 TIN으로부터 간략화한 것이다. 다시 말해, 아래의 TIN들은 하나의 지형을 정보량의 수(점의 수)에 따라 다양하게 표현한 것이다. [그림3]에서 볼 수 있듯이 TIN을 간략화하는 경우 전체적인 지형의 특징은 보존하고 있다. 그러나, 지형에서 유지해야 하는 중요한 형태적 정보마저 손실되는 경우가 발생한다. 즉, 그림 (가)와 (다)의 전체적인 모양은 유사하나 실제 고도 값을 추정하기 위해 계산단계에서 반드시 유지해야 하는 형태적 특성이 일부 소실됨으로 분석의 정확성을 떨어뜨린다.



실험에서 모델의 타당성을 평가하기 위해 사용한 기준은 질의의 정확성이다. 사용한 질의는 500개의 임의로 생성한 것으로 2차원 좌표에서의 3차원 값, 즉 주어진 2차원 점에서의 고도를 계산하는 것이다. TIN을 이용할 때 주어진 점에서의 고도를 계산하는 방법과 동일하게 제안한 본 논문에서 제시한 다각형에서의 높이 추정도 동일한 방법을 이용한다. 즉, 형태적 속성을 보존하고 있는 다각형에서 질의 점을 포함하고 있는 세 점을 구한다. 이 세 점은 삼각형을 이루게 되므로 TIN에서의 고도 계산과 같다. 아래 [그림 4]는 세 점을 또는 삼각형에서의 고도 값 계산을 표현한 소스 코드이다.

```

질의 점(2차원) : pt, 질의를 포함한 세 점(3차원) : p1, p2, p3
COORD w1, w2, w3, d; // 세 점을 이용해 추출한 상수 값
COORD elevation; // 추출하고자 하는 고도 값

d=p1.x*(p2.y-p3.y)+p2.x*(p3.y-p1.y)+p3.x*(p1.y-p2.y);
w1=(pt.x*(p2.y-p3.y)+p2.x*(p3.y-pt.y)+p3.x*(pt.y-p2.y))/d;
w2=(p1.x*(pt.y-p3.y)+pt.x*(p3.y-p1.y)+p3.x*(p1.y-pt.y))/d;
w3=(p1.x*(p2.y-pt.y)+p2.x*(pt.y-p1.y)+pt.x*(p1.y-p2.y))/d;
elevation=p1.z*w1+p2.z*w2+p3.z*w3;
    
```

[그림4] 세 점을 이용한 질의의 고도 값 계산 코드(C++)

실험에 사용된 척도는 다음과 같다. (나)의 TIN에서 계산된 고도 값을 정확하다고 가정하고 [그림2]의 다각형으로 구성된 형태적 특성을 고려한 지형 모델에서 계산한 값과 다각형과 유사한 점 개수를 가진 (다)의 TIN과 비교하는 실험이다.

아래의 [표2]는 이 실험의 결과이다. 형태적 특징을 고려한 지형 모델의 평균 오차율이 비슷한 데이터 크기인 TIN 보다 낮다. 이러한 결과는 형태적 특징을 고려한 모델을 이용하는 경우에는 최대한 지형의 형태적 특성을 반영함으로써 놓치지 말아야 할 정보를 유지하기 때문에 연산의 정확성을 높이기 때문이다. 이와 함께 제안한 모델은 평균 오차율 뿐 아니라 최대의 오차율도 낮은 장점을 보인다. [표2]를 보면 제안

한 모델은 TIN의 1/10의 정보량으로 약 3.7%의 오차를 가지게 된다. TIN과 제안한 모델이 유사한 크기의 데이터 양이라면 디스크에서 데이터를 적재하는데 드는 비용은 큰 차이가 없을 것이다. 이런 경우 오차율이 낮은 쪽이 성능이 뛰어나다고 말할 수 있다. 실험의 결과를 종합하면 제안한 모델은 적정 오차 수준으로 데이터 양을 크게 줄일 수 있을 뿐 아니라 비슷한 데이터 크기인 경우 TIN보다 나은 성능을 보여준다.

[표2] 점 567개의 TIN을 기준으로 한 분석 결과

	다각형	(다)
점의 개수(개)	240	256
평균 오차율 (%)	1.383	1.890
최대 오차율 (%)	23.323	87.595

5. 결론 및 향후 연구

본 모델의 가장 중요한 목적인 데이터 양을 줄이고 지형의 형태적 특징을 보존하자는 데 있다. 실험 결과에 따르면 형태적 특성을 고려한 지형 모델은 유사한 데이터 크기의 TIN에 비해 높은 수준의 정확성을 지니고 있음을 알 수 있다. 본 모델의 가장 중요한 목적인 데이터 양의 감소면을 살펴보면 [표2]의 경우 90% 가량의 데이터 양이 감소했다. 적정 수준의 오차는 허용 가능하다고 하면, 분석 시 실제 모델에서 유지하는 데이터 양의 감소로 인해 디스크에서 데이터를 참조하는데 걸리는 시간을 줄임으로써 성능을 향상시킬 수 있다. 동시에 형태적 특성을 보존함으로써 응용 프로그램에서의 분석의 용이함을 가져올 수 있다.

지형 데이터와 관련된 응용 레벨에서의 분석 내용은 상당히 다양하다. 실제 응용에 따라 데이터의 연산자와 색인 방법을 달리 해야 할 정도로 다양한데 현재 이 모델에서의 적용 가능한 연산자나 색인에 대한 연구가 없는 실정이다. 본 논문에서는 형태적 특성을 지닌 지형 데이터를 이용한 고도 추정과 관련된 부분의 실험만 수행하였다. 추후 이 부분에 관한 추가적인 연구가 이루어지게 되면 분석 범위의 한계도 극복될 수 있을 것이고 따라서 보다 다양한 응용 레벨에서의 모델에 대한 타당성 검증이 이루어질 수 있다.

6. 참고 문헌

- [1] M. van Kreveld, 'Digital elevation models and TIN algorithms', Springer-Verlag, LNCS 1997
- [2] L. P. Chew, 'Constrained Delaunay triangulations' Proc. Computational Geometry ACM 1987
- [3] A. Orocke, 'Computational Geometry in C', 2nd Edition, 1998
- [4] L. Floriani, 'Visibility algorithm on Triangulated Digital Terrain Models', Proc. IJGIS 1994
- [5] M. Kreveld, 'Trekking in the Alps without freezing or getting tired', Proc. ESA 1993
- [6] M. Kreveld, 'Drainage queries in TINs: from local to global and back again', Proc. SDH 1996
- [7] GIS on Massive Datasets
http://www.cs.duke.edu/geo*