

# 프랙탈 분석을 통한 비용효과적인 기준자료추출 알고리즘에 관한 연구

## A Cost Effective Reference Data Sampling Algorithm Using Fractal Analysis

\*김창재, \*\*이병길, \*\*\*김용일

\*서울대학교 지구환경시스템공학부 공간정보연구실 석사

\*\*서울대학교 지구환경시스템공학부 공간정보연구실 박사과정

\*\*\*서울대학교 지구환경시스템공학부 조교수

### 요 약

분류기법을 통해 얻어진 원격탐사 자료는 사용되기 이전에 그 정확성에 관한 신뢰도 검증을 해야 한다. 분류 정확도를 평가하기 위해서는 오분류행렬(confusion matrix)을 사용하여 정확도 평가를 하게 되는데, 이때 오분류행렬을 구성하기 위해서는 기준자료(reference data)에 대한 표본추출이 이루어져야 한다. 기준자료의 표본을 추출하는 기법간의 비교 및 표본 크기를 줄이고자 하는 연구는 많이 이루어져 왔으나, 추출된 표본들간의 거리를 줄임으로써 정확도 평가 비용을 감소시키고자 하는 연구는 미미한 실정이다. 따라서, 본 연구에서는 프랙탈 분석을 통하여 기준자료의 표본을 추출하였으며, 이를 바탕으로 기존의 표본추출 기법과 정확도 차이 및 비용효과 측면을 비교 분석하였다. 연구 결과, 프랙탈 분석을 통하여 표본을 추출하는 기법은 그 정확도 추정에 있어 기존의 표본추출 기법과 큰 차이를 보이지 않았으며, 추출된 화소들이 가까운 거리에 군집해 있어 비용효과측면에서 보다 유리함을 확인하였다.

### 1. 서 론

위성영상의 분류결과를 사용하기 위해서는 정확도 평가가 반드시 필요하며, 이를 위해서는 기준자료(reference data)의 추출이 요구된다. 이 작업은 상당한 시간과 노력이 소모된다. 기준자료의 표본을 추출함에 있어, 표본들간의 거리를 줄여 표본추출 비용을 감소하고자 하는 연구는 Todd(1980)와 Rhode(1978)에 의한 집락 추출법의 경제성 분석이 있었으나, 이 집락 추출법도 집락 내의 표본 크기가 크지 않고 집락들이 영상 전지역에 산재하여 있으므로 비용 면에서 크게 효과적이지 않다. 본 연구에서는 전체 영상을 대표하는 밀집된 표본을 얻는 방법으로 영상 전체에 대한 공간적 분포를 대표하는 국소지역으로부터 표본을 추출하는 방법을 제안하였으며, 영상 전체를 대표하는 국소지역을 찾아내기 위한 지표로는 자연현상의 한 특징인 자기 유사성(self-similarity)을 잘 정량화할 수 있는 프랙탈 차원을 사용하였다.

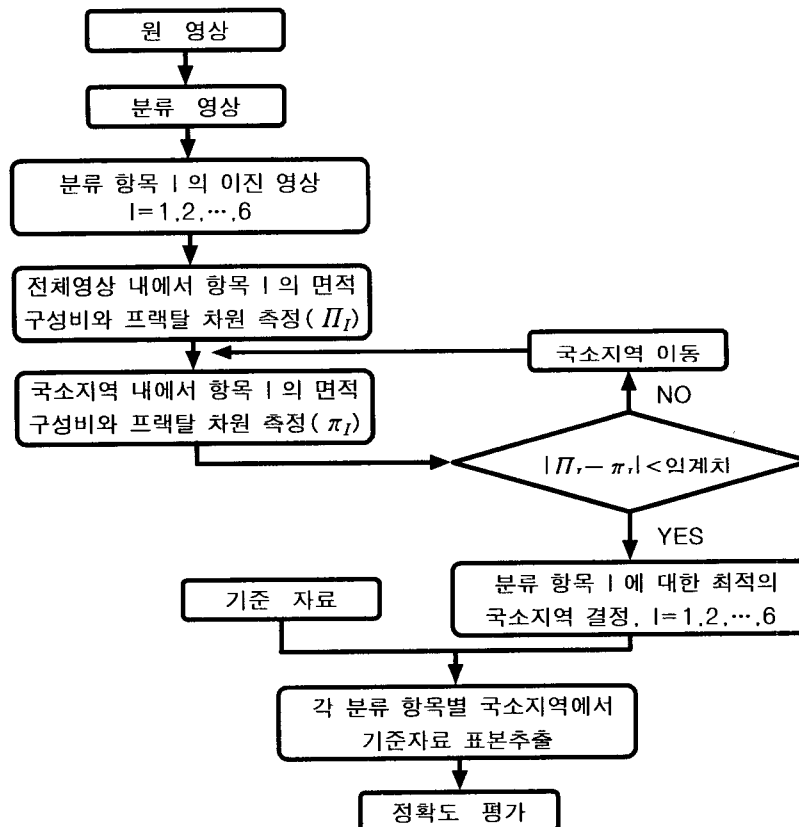
## 2. 프랙탈

유클리드 기하학과는 달리 프랙탈 기하학에서는 곡선과 곡면의 복잡한 정도에 따라 낮은 차원에서 높은 차원까지 갖는다. 공간적 복잡성(complexity)이나 변이성(variability)을 잘 정량화할 수 있다는 특성은 프랙탈의 커다란 장점이다. 또한 자기 유사성(self similarity)은 프랙탈의 또 다른 중요한 특성 중의 하나이다. 프랙탈 공간에서는 모든 축척에 대하여 항상 '전체'의 공간적 특성을 갖는 '부분'이 존재하게 된다. 이는 프랙탈 차원이 축척에 대해 독립적이라는 의미이며, 모든 축척에 대해 일정하다는 의미도 함축하고 있다. 자기 유사성은 '전체'와 '부분'간의 관련성이므로 축척에 따라 계급적으로 정렬되는 특성을 갖는다. 본 연구에서는 분류 영상의 프랙탈 차원을 계산하기 위해 isarithm 기법을 적절히 변형하여 사용하였다.

## 3. 실험

### 영상자료

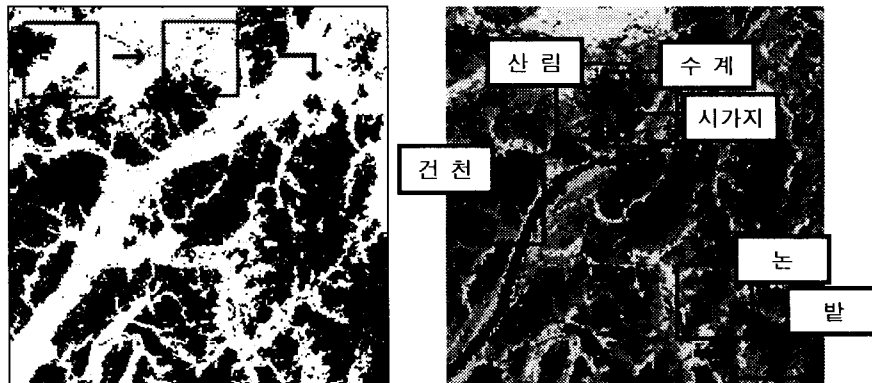
본 연구에 사용한 원영상은 1994년 5월에 취득된 공주, 청양 부근의 6개 Landsat TM 영상이며, 크기는 동일하게  $250 \times 250$ 이다. 분류 항목은 산림, 수계, 밭, 논, 건천, 시가지의 6개로 구성하였고 분류기법은 일반적인 최대우도분류기법을 사용하였다.



<그림 1> 프랙탈 분석을 통한 기준자료 표본 추출 알고리즘

### 프랙탈 분석을 통한 기준자료 표본추출 및 정확도 평가

기준자료의 표본 위치에 관한 정보를 추출하기 위해 우선 분류된 영상으로부터 각 분류 항목별로 이진 영상을 생성한 후, 각 이진 영상에서 영상 전체에 대해 해당 분류 항목이 차지하고 있는 면적 구성비와 프랙탈 차원을 계산한다. 다음으로 국소 지역의 크기를 결정하고 영상 전체에서 구한 것과 마찬가지로 국소 지역내에서 해당 분류 항목이 차지하고 있는 면적 구성비와 프랙탈 차원을 계산한다. 국소 지역내에서 계산된 값과 분류 영상 전체에서 계산된 값의 차가 일정 임계치 이하가 되면 그 국소 지역을 기준자료 표본추출 지역으로 선정한다. 이런 과정을 각 항목별 이진 영상에 대하여 국소 지역의 면적을 125×125, 100×100, 75×75, 그리고 50×50 등으로 점차 줄여 가면서 계산하고, 마찬가지로 각각에 대하여 최적의 국소 지역을 결정한다. 이렇게 결정된 국소 지역들의 위치로부터 이미 작성된 기준자료를 이용하여 정확도 평가를 위한 표본을 추출하게 된다. 프랙탈 기법, 단순 무작위 추출법, 계통적 추출법간에 비교하기 위해 표본의 크기는 프랙탈 기법에서의 국소 지역내 화소수와 모두 동일하게 추출한 후 정확도를 평가하였다. 정확도평가는 보편적으로 많이 사용하고 있는 오차행렬방법을 사용하였다.(그림 1, 2참조)

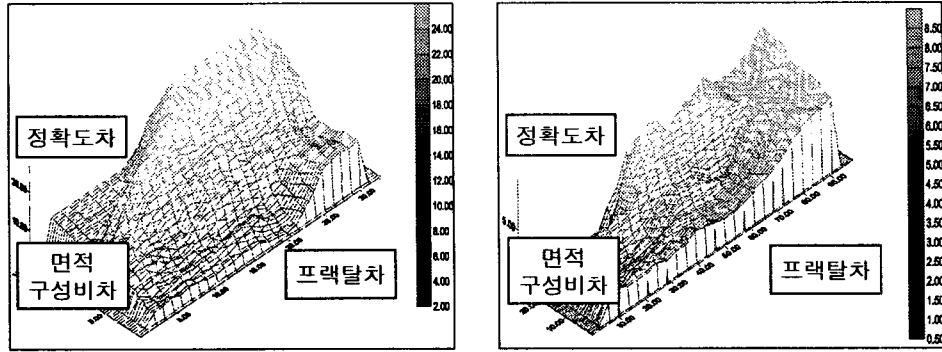


<그림 2> 산림의 이진 영상 및 국소 지역 이동(左)과 분류 항목별로 선정된 국소 지역(50×50)(右)

## 4. 결 과

### 프랙탈과 정확도간의 관계분석

연구 대상지 전지역에 대한 분류 영상의 프랙탈 차원과 각 항목별 면적 구성비를 구한 후, 국소 지역을 이동해 가며 계산한 프랙탈 차원과 각 항목의 면적 구성비로부터 전지역과 국소 지역간의 각 값의 차를 구한다. 또한 항공사진으로부터 대상지 전체에 대한 기준자료를 취득하고 있으므로 각 분류 항목에 대해 전체지역과 국소 지역간의 사용자 정확도 차이를 얻을 수 있다. 이들 세 변량, 즉 전체 영상과 이동하는 국소 지역간의 프랙탈차, 면적 구성비차, 정확도차를 3차원의 그래프로 도식하면 그림 3과 같다. 이 그림에서 보듯이 국소 지역의 프랙탈 차원과 면적 구성비가 전체 영상의 프랙탈 차원과 면적 구성비에서 멀어질수록 각 항목별 정확도의 차이가 커짐을 알 수 있다.



<그림 3> 논(左)과 시가지(右)의 프랙탈차, 면적구성비차, 정확도차간의 관계 그래프 (국소 지역 : 125×125)

### 분류 정확도 평가 결과

분류 영상 전체의 프랙탈 차원 및 각 항목별 면적 구성비를 구하고, 국소 지역을 이동해 가며 이들 값의 차이를 가장 적게 보이는 국소 지역을 선정한 후, 그 지역들에서 기준자료의 표본을 추출하였다. 이 추출된 표본들로 오차행렬을 구성하고 전체 정확도를 평가하였다. 이 결과와, 기존의 표본추출 기법인 단순 무작위 추출법과 계통적 추출법으로부터 추출한 표본으로 정확도를 평가한 결과를 비교 분석하였다.

표 1은 기준자료 표본추출 기법별로 정확도를 평가한 후 이 기법들간의 정확도 차이를 나타낸 것이다. 정확도의 차이는 절대값을 취하였다.

표 1의 결과를 바탕으로 각 기법에 대하여 국소 지역의 크기별 정확도 차이의 평균에 관해 표 2와 같이 나타내었다. 표 2를 보면 전수 추출과 프랙탈 추출에 의한 정확도 평가의 차이가 1.11-1.99% 사이에 존재함을 알 수 있다. 즉 프랙탈 기법으로 기준자료를 표본추출한 후 정확도를 평가하였을 때, 전수 조사와의 정확도 차이가 약 2% 안쪽임을 보이고 있다. 기존의 표본추출 기법들과 비교해 보아도 무작위 추출법과는 그 차이가 1.33% 안쪽이고, 계통적 추출법과는 차이는 2.17% 안쪽으로 근소한 차이를 보이고 있다.

### 표본추출 면적에 관한 분석

표본의 추출면적은 기존의 단순 무작위 추출법과 계통적 추출법이 본 연구의 대상지 총면적  $56 \text{ km}^2$  내에서 표본을 추출해야 함에 반해, 본 연구에서 제안한 프랙탈 기법을 통하면 국소지역 크기가  $125 \times 125$ 일 때는  $14 \text{ km}^2$ ,  $100 \times 100$ 일 때는  $9 \text{ km}^2$ ,  $75 \times 75$ 일 때는  $5 \text{ km}^2$ , 그리고  $50 \times 50$ 일 때는  $2.25 \text{ km}^2$  내에서 표본을 추출하면 된다. 즉, 총 대상지에 대해 25%~4%의 면적만으로 기존의 추출기법을 통한 정확도와 근사한 결과를 산출하는 표본을 추출할 수 있다.

<표 1> 기준자료 표본추출 기법간 추정된 정확도 차이(%)

국소지역 크기	비교 대상	전수 추출과	전수 추출과	전수 추출과	무작위 추출과	계통추출과
		무작위 추출	계통 추출	프랙탈 추출	프랙탈 추출	프랙탈 추출
영상 1	125×125	0.26	0.02	0.06	0.27	0.08
	100×100	0.34	0.16	2.37	0.24	2.53
	75×75	0.35	0.12	0.33	0.43	0.45
	50×50	0.80	0.15	2.88	0.28	2.73
영상 2	125×125	0.37	0.35	0.98	1.09	1.33
	100×100	0.36	0.03	0.75	0.77	0.77
	75×75	0.52	0.57	1.90	1.77	2.46
	50×50	0.76	0.37	1.16	1.23	0.79
영상 3	125×125	0.22	0.14	1.64	1.66	1.50
	100×100	0.29	0.00	0.98	1.06	0.98
	75×75	0.48	0.10	0.15	0.45	0.06
영상 4	125×125	0.32	0.19	1.68	1.57	1.87
	100×100	0.29	0.19	2.88	2.78	3.07
	75×75	0.55	0.29	1.48	1.45	1.77
영상 5	125×125	0.19	0.13	1.35	1.40	1.22
	100×100	0.33	0.37	1.15	1.27	0.78
	75×75	0.48	0.27	1.89	1.95	2.16
	50×50	0.73	0.85	1.64	1.52	2.50
영상 6	125×125	0.25	0.21	0.97	0.86	1.18
	100×100	0.37	0.04	0.94	1.03	0.91
	75×75	0.46	0.44	1.47	1.62	1.03
	50×50	0.63	0.35	2.31	2.29	2.66

<표 2> 분류 영상 6개에 대한 각 기법간 정확도 차이의 평균(%)

국소지역 크기	비교 대상	전수 추출과	전수 추출과	전수 추출과	무작위 추출과	계통추출과
		무작위 추출	계통 추출	프랙탈 추출	프랙탈 추출	프랙탈 추출
	125×125	0.27	0.17	1.11	1.14	1.20
	100×100	0.33	0.13	1.51	1.19	1.51
	75×75	0.47	0.30	1.20	1.28	1.32
	50×50	0.73	0.43	1.99	1.33	2.17
	평 균	0.45	0.26	1.46	1.24	1.55

### 5. 결 론

본 연구에서 제시한 프랙탈 분석을 이용하여 표본을 추출한 후 정확도평가를 한 결과, 정확도 추정과 비용면에서 다음과 같은 결론을 도출하였다.

첫째, 프랙탈 분석을 통한 표본추출 기법을 사용하여 표본을 추출한 후 정확도를 평가해 본 결과, 본 연구의 실험대상지역의 경우 전수 조사에 의한 정확도 평가 결과와 2% 이내에서 차이를 보였다. 또한 국소 지역의 각 크기에 따른 전수 조사와의 정확도 차이를 평균한 결과는 1.46%였다. 따라서 본 연구에서 제시한 표본추출 기법의 정확도 추정 성능이 크게 떨어지지 않음을 확인하였다.

둘째, 기존의 표본추출 기법과의 정확도 차이를 비교한 결과, 무작위 추출과는 평균 1.24%의 차이를, 계통적 추출법과는 1.55%의 차이를 보였다. 이 결과로 기존의 표본추출 기법과, 정확도 추정에 있어 큰 차이를 보이지 않음을 확인하였다.

셋째, 프랙탈 분석을 통하면 영상 전지역의 정확도를 대표할 수 있는 국소 지역 추출이 가능함을 확인하였다.

넷째, 정확도 평가에 사용되는 기준자료의 표본을 추출함에 있어, 프랙탈 분석기법을 이용할 경우 기존의 기법에 비해 국소지역의 크기에 따라 25%~4%정도의 면적만 관측하면 된다. 따라서 본 연구에서 제시한 프랙탈 기법을 사용하여 표본을 추출할 경우, 국소지역의 크기가 점차 감소함에 따라 기존의 표본추출기법들보다 비용의 감소측면에서 매우 효과적임을 확인할 수 있었다.

#### 참고 문헌

- 1) 남영우, 계량지리학, 법문사, 1995
- 2) 어양담, 위성영상 분석을 위한 트레이닝 정규화 알고리즘과 클래스 분리도 측정기법 개발, 서울대학교 대학원 공학박사학위논문, 1999
- 3) Batty, Michael and Longley, Paul, *Fractal Cities*, Academic Press, 1994
- 4) Clarke, Keith C. and Schweizer, Diane M., Measuring the Fractal Dimension of Natural Surfaces Using a Robust Fractal Estimator, *Cartography and Geographic Information Systems*, Vol. 18, No. 1, pp. 37-47, 1991
- 5) Congalton, Russell G., A Comparison of Sampling Schemes Used in Generating Error Matrices for Assessing the Accuracy of Maps Generated from Remotely Sensed Data, *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 5, pp. 593-600, 1988
- 6) Congalton, R.G., A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data, *Remote Sensing of Environment*, 37, 35-46, 1991,
- 7) De Cola, Lee, Fractal Analysis of a Classified Landsat Scene, *Photogrammetric Engineering and Remote Sensing*, Vol. 55, No. 5, pp. 601-610, 1989
- 8) Goodchild, Michael F., Fractals and the Accuracy of Geographical Measures, *Mathematical Geology*, Vol. 12, No. 2, pp. 85-97, 1980
- 9) Jaggi, S., Quattrochi, Dale A., and Lam, Nina Siu-Ngan, Implementation and Operation of Three Fractal Measurement Algorithms for Analysis of Remote-Sensing Data, *Computers & Geosciences*, Vol. 19, No. 6, pp. 745-767, 1993
- 10) Lam, Nina Siu-Ngan and De Cola, Lee, *Fractals in Geography*, PTR Prentice-Hall, 1993
- 11) Qiu, Hong-lie, Lam, Nina Siu-Ngan, Quattrochi, Dale A., and Gamon, John A., Fractal Characterization of Hyperspectral Imagery, *Photogrammetric Engineering and Remote Sensing*, Vol. 65, No. 1, pp. 63-71, 1999
- 12) Story, M. and Congalton, R. G., Accuracy Assessment : A User's Perspective, *Photogrammetric Engineering and Remote Sensing*, Vol. 52, No. 3, pp.397-399, 1986