한국측지학회지, 제 17 권 제 2 호, pp. 129~136(1999년 6월 30일)

공간분포지표를 이용한 위성영상 분류오차의 공간적 분포 평가
Estimating the Spatial Distribution of Satellite Image Classification Error Using Index of Spatial Distribution

이병길* · 김용일** · 어양담***
Lee, Byoung-Kil · Kim, Yong-Il · Eo, Yang-Dam

要 旨

영상분류 결과는 지형적 영향, 영상의 상태 등에 따라 전체 영상에 대하여 균일하지 않을 수 있다. 본 연구에서는 분류 결과의 불균일성과 위성영상 분류 오차의 공간적 분포를 평가하기 위해 ISDd (Index of Spatial Distribution by distance) 와 ISDs (ISD by scatteredness)의 개념을 제안하였다. ISDd는 지표화된 오분류 화소 간의 거리이고, ISDs는 오분류 화소의 산포도에 관한 통계적 지표이다. 실제 위성영상에 대한 실험을 통하여 ISDd와 ISDs를 계산 및 평가하였으며, 실제 국지적 오분류 영역을 추출하여 오분류의 원인을 고찰하였다. 본 연구 결과, ISDd와 ISDs를 동시에 사용하여 오분류 화소의 국지적 밀집 여부와 밀집 정도의 평가가 가능하였으며, 그 결과를 토대로 영상의 일부분에 대한 분류결과의 채택/기각을 결정할 수 있었다. 따라서, 전체 분류정확도 외에 공간분포지표를 사용함으로써 사용자는 오분류 화소의 공간적 분포 상태를 파악할 수 있으며, 분류 결과의 적합성 및 신뢰성 판단을 위한 추가적인 기준을 가질 수 있다.

ABSTRACT

The quality of image classification results is not always uniform over entire image. Thus, this study proposes the concept of ISDd (Index of Spatial Distribution by distance) and ISDs (ISD by scatteredness) for the evaluation of unevenness of result quality, and spatial distribution of satellite image classification errors. The ISDd is indexed mean distance of misclassified pixels and the ISDs is statistical indicator of scatteredness of misclassified pixels. In this study, the ISDd and the ISDs are calculated and evaluated for some satellite images, then misclassified area is extracted and the reasons of misclassification are examined. As the result of this study, using both the ISDd and the ISDs, the basis of decision on adoption/rejection of classification results is offered at sub-image level by evaluation of the local aggregation of misclassified pixels. Using Index of Spatial Distribution, as well as overall classification accuracy, users can understand the spatial distribution of misclassified pixels, and can have the additional criterion of the judgement on suitability and reliability of classification results.

1. 서 론

위성영상이 일반에게 공개된 1972년 이후, 원격탐사분야의 대부분의 연구는 영상 데이터로부터 정확한 정보를 추출하는데 초점을 맞추고 있다.[1)]

원격탐사는 크게 두 가지 분야로 대별되는데, 그 중 하나는 경사관측이 가능한 위성의 입체 영상을 이용하여 DEM을 구축하는 것과 같은 정량적인 분야이고, 나머지 하나는 화소의 수치값(DN ; Digital Number)을 통계적으로 분석하여 공간적 정보를 얻는 정성적 분야이다.

최근까지도 주요 연구 주제 중의 하나는 위성 데이터 분석의 정확도 향상에 관한 것이었다. 그 결과 영상으로부터 도시지역의 복잡한 공간분포를 분석하는 것도 가능해졌다.[2)] 특히, 분류 방법의 개발을 통한 분류정확도의 향상은 원격탐사 기술의 실용화에 매우 중요한 역할을 담당하였다.

이와 같은 이유로 분류 결과에 대한 객관적이고 정확한 평가방법에 대한 연구도 광범위하게 수행되어 왔다.

*서울대학교 도시공학과 박사과정
**서울대학교 지구환경시스템공학부 조교수
***서울대학교 공학연구소 특별연구원

Congalton과 Mead(1983)는 분석자 간의 유사성과 분석의 변수 간의 유사성을 Kappa 계수와 정규화 된 Z-통계치를 이용하여 정량적으로 평가하였다.[4]

Rosenfeld와 Fitzpatrick-Lins(1986)는 분류 정확도의 추정에 Kappa 계수를 사용하여 전체 정확도와 분류항목별 정확도를 추정하였다.[6]

이러한 연구는 영상의 분류 결과에 대한 전체 정확도의 정량화 방법을 제공하였으나, 연구 대상지역의 지형, 토지 피복의 국지적 변동과 영상 자체의 잡음 등에 의해 분류 결과의 품질은 전체 영상에 대하여 균일할 수 없다. 이러한 영향은 그림자 등에 의해서는 면형의 오차를, 잡음에 의해서는 선형 또는 점형 오차를, 소단위 토지피복에 의한 혼합화소(mixel : mixed pixel)에 의해서는 군집된 점형의 오차를 나타낼 수 있을 것이다. 이러한 오차의 영향이 크다면, 전체 정확도가 지도제작의 목적에 적합하더라도 부정확하게 분류된 지역의 분류 결과는 sub-image 수준에서 기각되어야만 하는 경우가 발생할 수도 있고, 전체 정확도는 기각역에 포함될지라도 부정확하게 분류된 지역이 제거되거나 재분류 된다면, 영상의 나머지 부분은 사용 가능한 경우가 발생할 수도 있다.

본 연구에서는 오차 행렬을 사용하는 기존의 분류 정확도 평가기법을 살펴보고, 위성영상 분류 오차의 공간적 분포의 평가를 위해 ISDd와 ISDs의 개념을 제안하였다. 항목별 분류 정확도는 전체 화소 수에 대한 정확하게 분류된 화소 수의 비율로 오분류 화소의 공간적인 분포에 대한 정보는 전혀 포함하고 있지 않다. 따라서, 분류 품질의 불균일성을 설명할 수 있는 지표를 도입함으로써, 분류 결과의 신뢰성과 유용성을 증진시킬 수 있다.

이러한 목적을 달성하기 위해, 첫째 점분포의 공간적 유형을 판정하는 기존의 기법을 살펴보고, 둘째 분류 오차의 공간적 분포를 평가하는 방법을 연구하여 제안하고, 마지막으로 항목별 분류 정확도의 추정치와 분류 오차의 공간분포에 대한 지표를 시범적으로 적용해 보았다.

2. 점분포 유형의 판정

점분포 유형을 판정하기 위한 기존의 분석 방법은 주로 환경학자에 의해 제안되었으며, 환경학자와 지리학자에 의해 발전되었다. 이러한 방법은 크게 방격법과 거리법의 두 가지로 분류된다.

2.1 방격법

방격법에서 연구 대상지역은 균일한 크기의 격자로 분할되고, 각 격자에 포함되는 점의 수가 계산된다. 이 때, 각 점이 임의로 분포하고 있다면, 점의 분포는 포아송(Poisson) 분포를 따르게 된다. 격자의 수(k)와 점의 수(N)가 주어졌을 때, 하나의 격자에 하나의 점이 포함될 확률은 $p = \lambda/N$이다. 여기서, λ는 점의 밀도이고, p는 포아송 분포, $e^{-\lambda}\lambda^m/m!$를 따른다. 이 때, m은 단위면적에 포함된 점의 수를 말한다. 따라서, 포아송 분포의 특성에 따라 점분포의 평균(λ)은 분산(s^2)과 같아야 한다.

이러한 분포의 통계적 검정에는 χ^2-test가 사용되며, 검정통계량은 $\chi^2_{k-1} = (k-1)\ (s^2/\lambda)$이다.

주어진 점분포에 대해 방격법을 적용할 때는 격자의 크기가 분석 결과에 큰 영향을 미친다. 격자의 크기가 증가하면 통계치는 연구 대상지역에 밀집된 분포가 존재함을 보이는 경향을 가지게 된다. 따라서, 격자의 크기는 방격법의 적용에 가장 중요한 요소로 작용하며, λ에 대해 Bartlett, Greig-Smith 및 Curtis와 McIntosh는 각각 1.6, 1.0 and 4.0이 적합하다고 제안하였다.[8]

방격법은 격자 내에서의 산포에 대해서는 설명력이 좋으나, 격자 간의 산포에 대해서는 그렇지 못하다.[3] 그림 1에서와 같이, (a)와 (b)가 동일한 통계치를 가질 때, (a)가 (b)에 비해 더 군집화 된 유형을 가지고 있을 수 있다.

2.2 거리법

거리법에서는 $\bar{r}_e$에 대한 $\bar{r}_a$의 비율을 이용한다. 여기서, $\bar{r}_e$는 임의로 분포된 점의 쌍이 갖는 최소거리 평균의 기대값이고, $\bar{r}_a$는 각 점으로부터 계산된 최근린 점까지의 평균거리이다. 두 점 (U_i, V_i)과 (U_j, V_j)간의 거리를 유클리드 거리라고 한다면, n개 점 간의 평균최소거리는 다음과 같이 주어진다.

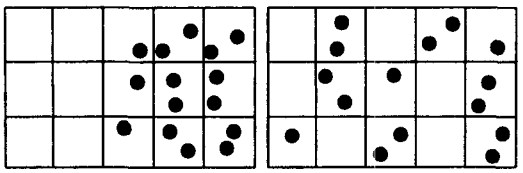


그림 1. 동일한 결과의 상이한 점분포

$$\bar{r}_a = \frac{\sum_i \min_j \sqrt{(U_i - U_j)^2 + (V_i - V_j)^2}}{n} \quad (1)$$

또한 최소거리 평균의 기대값은 $\bar{r}_e = 1/2\sqrt{\lambda}$ 이다. 여기서, λ는 단위면적 당 점밀도의 기대값이다.

만약 점들이 임의적으로 분포되어 있다면, $\bar{r}_a$는 $\bar{r}_e$와 같아야만 한다. 이의 통계적 검정에는 Z-검정이 사용되고, 검정통계량은 $Z = (\bar{r}_a - \bar{r}_e)/SE_d$이다. 여기서 $SE_d = 0.26236/\sqrt{n\lambda}$ 이다.

이 방법은 연구대상 지역이 확대될수록 밀집된 분포로 인식하는 편의를 가진다. 게다가, 경계선 부근의 점들은 다른 점에 비해 긴 최소거리를 갖게 된다. 이러한 문제를 피하기 위해, transect of quadrat, toroidal edge correction 그리고 ordered neighbor distance method 등의 기법이 제안되었으나, 연구대상지역의 크기와 경계선 부근의 점에 대한 문제는 여전히 해결되지 않고 있다.[8,3)]

3. 분류오차의 공간적 분포 평가

앞절에서 언급한 바 대로, 점분포 유형을 판정하는 기존의 방법은 몇가지 문제점을 가지고 있다. 게다가, 이러한 방법들은 유형 생성 과정의 임의성을 논증하기 위한 방법으로 주로 사용되어 왔다. 따라서, 기존의 방법으로는 점분포에 군집이 포함되어 있는지의 여부는 판정할 수 있으나, 군집된 유형을 분류하기는 어려웠다. 여기서, 군집이란 오분류된 화소가 근거리에 밀집되어 있는 상태를 말한다.

이와 같은 문제를 해결하기 위해, 본 연구에서는 ISDs와 ISDd의 개념을 제안하고, 이 두가지 지표를 동시에 사용하는 것을 제안하였다. 그 결과로, 영상의 크기나 오분류된 화소수에 독립적이며, 군집의 유형을 설명할 수 있는 지표를 얻을 수 있게 되었다.

3.1 분류오차의 공간적 분포 측정

분류오차의 공간적 분포를 측정하기 위해 제안된 ISDd는 지표화된 오분류 화소 간의 거리이고, ISDs는 오분류 화소의 산포도에 대한 통계적 지표이다.

3.1.1 공간분포 거리지표(ISDd : Index of Spatial Distribution by distance)

ISDd 계산을 위한 첫번째 단계로, 오분류 화소 간의 거리 합을 거리의 조합수로 나눈 평균거리를 계산해야 한다(식 2).

$$\text{평균거리} = \frac{l_{12} + l_{13} + \cdots + l_{1n} + l_{23} + \cdots + l_{n-1n}}{{}_nC_2} \quad (2)$$

여기서,

$l_{12} = \overline{p_1 p_2}$

$\vdots \qquad \ddots$

$l_{1n} = \overline{p_1 p_n} \quad l_{2n} = \overline{p_2 p_n} \quad \cdots \quad l_{n\text{-}1n} = \overline{p_{n-1} p_n}$

n : 오분류 화소 수

p_i : i번째 오분류 화소

위성 영상 분류오차의 공간적 분포 판정을 위한 오분류 화소 간의 평균 거리는 영상의 크기에 따라 달라진다. 큰 영상의 평균 거리가 작은 영상의 평균 거리에 비해 커지기 때문에, 크기가 다른 두개 이상의 영상의 분류 결과를 객관적으로 비교하는 것이 어렵게 된다. 그래서, 평균거리는 식 (3)과 같은 방법에 의해 지표화되어야 한다. 여기서, ISDd*는 평균 거리를 영상의 대표거리로 나눈 값이다. 본 연구에서는 대표거리를 영상의 높이와 넓이의 평균으로 정의하였다.

$$ISDd^* = \frac{\text{평균거리}}{(V+H)/2} \quad (3)$$

여기서,

V : 영상의 높이

H : 영상의 넓이

수회에 걸친 모의 실험을 통해, 오분류 화소의 수, 화소의 물리적 규격및 오분류 화소의 분포 유형에 따라 ISDd*가 변화됨을 알 수 있었다.

1) 오분류 화소의 수

오분류 화소의 절대적인 수에 비해 전체 화소에 대한 오분류 화소의 비율인 상대적인 수가 중요한 변수가 된다. m×n의 크기를 가진 영상에 P개의 오분류 화소가 포함되어 있다고 가정하자. 만약 P/(m×n)이 0에 가깝다면, ISDd*는 0에서부터 ISDd*의 최대값까지 가질 수 있는데, 이러한 경우에는 분류가 매우 정확히 된 것으로 오분류 화소의 공간적 분포가 큰 의미를 가지지 않는다. 만약, P/(m×n)가 1에 가까워 진다면, ISDd*의 값은 규칙적인 배열에 대한 ISDd*의 값을 갖게 되며, 이 경우에 분류 결과는 오분류 화소의 분포와 무관하게 기각되

어야 한다.

ISDd*의 이론적인 최대값은 오분류 화소가 동일한 위치에 중복해 있는 것을 허용하면 얻을 수 있다. 오분류 화소의 1/4 만큼이 영상의 구석점(좌상, 우상, 좌하, 우하)에 중복하여 위치한다면, 단위 높이와 넓이를 가진 영상에 대해 ISDd*는 다음과 같이 계산된다.

$$ISDd^* = \frac{(\sqrt{2}\times n1 + 1\times n2 + 0\times n3)}{(n1+n2+n3)} \quad (4)$$

여기서,

n1 : 거리가 $\sqrt{2}$ 인 화소 쌍의 수

n2 : 거리가 1인 화소 쌍의 수

n3 : 거리가 0인 화소 쌍의 수

오분류 화소의 수가 4n인 경우, ISDd*의 최대값은 식(5)과 같이 유도된다.

$$ISDd^* = \frac{(\sqrt{2}+2)n^2}{4n^2-n} \quad (5)$$

식에서 알 수 있는 것과 같이 ISDd*는 오분류 화소의 수가 증가함에 따라 감소하는 경향을 보인다. ISDd* 최대값의 극한값은 0.85355이다.

n×n 규칙 배열된 오분류 화소에 대한 ISDd*의 값은 orthogonal 방향의 거리의 합과 diagonal 거리의 합으로 구분하여 계산할 수 있다. 수치해석법에 의해 계산된 결과는 39×39(즉, 1521점) 이상에서 1/10,000이하의 차이를 보이며, 0.5215정도의 값에 수렴한다.

그러나, 실제의 위성영상에서는 오분류 화소의 중복이 발생하지 않는다. 따라서, 오분류 화소의 수는 전체 화소 수를 넘어설 수 없다.

2) 화소의 물리적 규격

이론적인 두 점간의 최대 거리는 영상의 최대 높이와 최대 넓이를 가지고 계산될 수 있으나, 실제 위성영상의 경우, 두 화소 간의 거리는 각 화소의 중심 좌표를 이용하여 계산된다. 그런데, 영상의 높이와 넓이는 화소의 물리적 규격에 화소의 행수와 열수를 각각 곱하여 얻어진다. 즉, 화소 간의 거리는 일반적인 점의 거리에 비해 화소 1개만큼의 이격을 가지게 된다. 이러한 이유로 식(3)에 사용된 대표거리가 보정되어야 하고, 이에 따라 ISDd* 역시 보정되어야 한다.

식(6)는 영상의 4개 구석점에 오분류 화소 4개가 위치한 경우에 계산된 ISDd*이다.

$$ISD_d^* = \frac{1}{{}_4C_2}\left(1-2\cdot\frac{1}{2}\cdot\frac{1}{m}\right)(4+2\sqrt{2}) \quad (6)$$

ISDd*는 식(6)에서 보인 바와 같이 (1-1/m)에 따라 크기가 변화한다. ISDd*는 영상의 공간 해상력이 증가함에 따라(즉, 화소의 물리적 규격이 감소함에 따라) 감소한다. 이 때, m이 150보다 크다면, ISDd*의 변화량(감소분)은 0.01보다 작아지게 되어 화소의 물리적 규격에 의한 영향은 거의 무시될 수 있다. 그렇지 않은 경우에는 식(3)의 분모에 식(7)을 사용하여 화소의 물리적 규격에 의한 영향을 최소화할 수 있다.

$$(V' + H')/2 \quad (7)$$

여기서,

V'=V-V/m, H'=H-H/m

3) ISDd*의 정규화

앞에서 언급한대로 ISDd*는 오분류 화소 수(n)의 함수이다.

n=4일 때, ISDd*는 0에서 1.138까지의 값을 가지며, n>100일 때, 0에서 약 0.85까지의 값을 갖게 된다. ISDd*를 지표로 사용하기 위해서는 식(8)을 사용하여 정규화시켜 ISDd로 만들어야 한다.

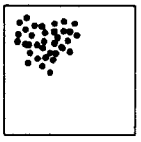
$$ISDd = 2.7 \cdot ISDd^* \cdot E^{-ISDd^*} \quad (8)$$

정규화의 결과로, ISDd는 0에서 1사이의 값을 갖게 되었고, 점분포 유형 만의 함수가 되었다.

4) 오분류 화소의 분포

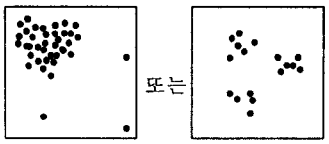
오분류 화소의 분포 유형은 다음의 그림과 같은 형태로 유형화 할 수 있다;

a. 0 < ISDd <= 0.54



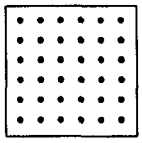
: 과밀한 군집 (즉, 하나의 사분면 내에 밀집)

b. 0.54 < ISDd <= 0.86

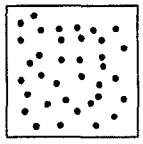


: 약간의 군집이 근거리에 위치 (즉, 둘이나 그 이상의 사분면에)

c. 0.86 < ISDd



: 규칙적으로 분포



: 임의적으로 분포

여기에서, 이와 같은 분포에 대한 ISDd를 구분하는 값들은 유형 각각에 대해 여러 번의 모의실험 결과를 통해 추출되었다.

3.1.2 공간분포 산포지표(ISDs : Index of Spatial Distribution by scatteredness)

오분류 화소의 산포도를 평가하기 위해서는 다음과 같은 가정이 필요하다.

만약 오분류가 우연오차에 의해 만들어진 것이라면, 오분류 화소가 영상의 일정 부분에 위치할 확률은 고정되어 있다. 만약 영상을 몇 개의 격자로 분할한다면, λ개의 오분류된 화소가 하나의 격자에 포함될 확률은 λ=P/k (k = 격자의 수, P = 오분류 화소의 수)인 포아송 분포를 따른다.

포아송 분포의 중요한 특성 중 하나는 평균과 분산이 동일하다는 것이다. 그러므로, 각 격자에 포함된 오분류 화소 수의 평균(λ)와 분산(υ)를 이용하여 산포도에 대한 지표를 구할 수 있다.

$$ISDs = \upsilon/\lambda \quad (9)$$

만약 오분류 화소가 임의적으로 분포하고 있다면, ISDs는 1이 된다. 만약 오분류 화소가 규칙적으로 분포하고 있다면, 각 격자에 포함된 오분류 화소의 수는 λ가 되므로 ISDs는 0이 된다. 그리고, 만약 오분류 화소가 밀집되어 분포하고 있다면, 각 격자에 포함된 오분류 화소의 수는 λ와 달라지게 되고 ISDs는 1보다 크게 된다.

2.1절에서 살펴본 바와 같이 방격법에서 분석결과에 가장 큰 영향을 주는 것은 방격의 크기이다. 본 연구에서는 이희연(1991)[1]에 따라 λ=1을 사용하여 격자의 크기를 결정하였다.

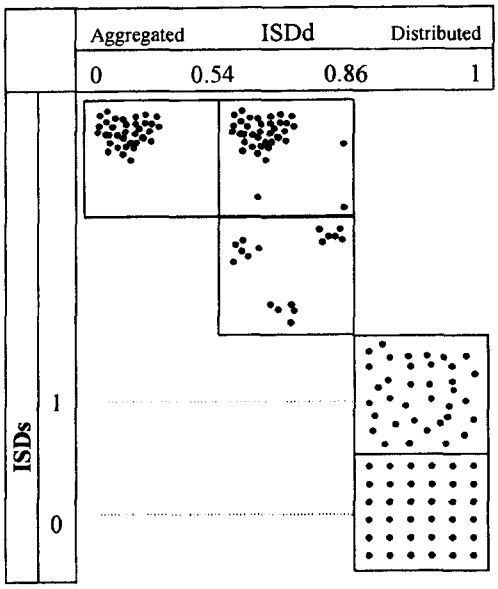


그림 2. ISDd/ISDs를 이용한 공간 분포 유형의 결정

3.2 공간분포 유형의 결정

ISDd와 ISDs는 상호보완적인 관계를 가진다. 따라서, 두 개의 지표를 동시에 사용하면, 공간분포 유형을 그림 2과 같이 결정할 수 있다.

3.3 오분류 영역의 추출

ISDd와 ISDs를 이용하여 분류 결과에 군집이 있다고 판단 되면, 전체정확도 또는 평균과 같은 일정한 기준과 비교하여 그 이상의 오분류 화소를 갖고 있는 영역을 추출할 수 있다.

주어진 정확도에 비해 낮은 전체 정확도를 갖는 window를 오분류 영역으로 선택하도록 하여, window를 전체 영상에 대해 이동시켜 원래의 분류 결과로부터 과도한 오분류 화소를 가지고 있는 오분류 영역을 추출한다.

Window의 크기는 통계의 정규성을 위하여 기준 데이터가 30개 이상 포함될 수 있는 크기로 하는 것이 좋다.[7]

추출된 오분류 영역은 사용된 분류방법에 의해서는 일정 수준 이상의 정확도를 확보할 수 없는 지역이므로, 분류에 사용된 것과는 다른 기법(예를 들어, 분류에 통계적 분류를 시행하였다면 인공신경망, 전문가 시스템 등과 같은 비모수적 분류 기법)을 이용하여 재분류하거나,

[1]공간적 경쟁이 심한 경우의 지리적 분석을 위해서는 A/N (A is area, N is number of points)이 방격법의 격자 크기 결정에 적합한 것으로 간주된다[3]. 이는 λ=1을 의미한다.

추출된 영역의 결과를 전체적으로 기각하여 사용하지 않는 방법을 취해야 한다.

4. ISDd/ISDs의 적용 및 평가

4.1 연구대상 지역및 사용 데이터

본 연구는 1:50,000 지형도에서 공주, 청양에 해당되는 지역을 연구대상으로 하였다. 대상 지역은 도시, 논, 밭, 산림, 하천 등 지형지물이 적절히 혼재되어 있어 원격탐사 영상분류와 그에 관련된 연구에 적합한 지역이다. 위성영상으로는 LANDSAT의 MSS, TM 및 SPOT의 XS를 사용하였고 기준자료는 1:50,000 지형도에서 구축하였다. 표본은 계통적인 추출법으로 1,200개를 1000 m 간격으로 추출하였고, 표본의 수가 적은 분류항목에 대해 42개의 표본을 랜덤 추출법으로 추출하였다.

사용된 위성영상은 MSS와 TM은 1994년 5월, SPOT XS는 1995년 4월에 취득되었다. 좌표등록을 위한 기하보정의 변환식으로는 affine변환식을 화소재배열 방법으로는 최근린내삽법을 사용하였으며, 이때, RMSE는 MSS, TM, XS 영상에 대해 각각 0.4, 0.5, 0.7화소였다. 분류에는 MLC가 사용되었으며, 사용된 밴드는 MSS는 4, 5, 7; TM은 2, 3, 4; XS는 1, 2, 3이 각각 사용되었다.

분류항목은 대상지역의 지형공간 분포를 고려하여 표 1과 같이 6개로 하였다.

각각의 영상에 대해서 MLC 기법을 사용하여 분류하고 전체 정확도, 정규화를 거친 Kappa 계수, ISDd와 ISDs를 구하여 분석을 수행하였다.

연구 수행에 필요한 모든 프로그램은 Borland C++ Builder(Borland사 제품)로 컴파일하였으며, 영상자료 분류를 위해 ER Mapper 5.2(Earth Resource Mapping Pty Ltd사 제품)를 사용하였다.

표 1. 클래스 별 기준 데이터 수

클래스 번호	클래스	기준 데이터 수
1	논	262
2	밭	73
3	주거지	76
4	호수	10
5	강	28
6	산림	793
계		1,242

4.2 연구대상지역 적용 결과 분석

분류결과의 전체 정확도, Kappa계수 및 ISDd, ISDs는 표 2와 같이 계산되었다.

표 2에서 알 수 있는 것처럼 항목별 분류정확도는 영상에 따라 변동을 보이는 반면에 ISDd/ISDs는 전체적으로 안정된 값을 보인다. 이는 ISDd/ISDs가 연구 대상지역의 공간적 또는 영상 자체의 특성에 기인한 대상지역 전체에 대한 오분류 화소의 분포와 발생 가능성을 보여주는 값이기 때문이다.

MSS와 TM은 1.0보다 약간 작은 ISDs값을 가진다. 이는 규칙성을 가진 임의 분포임을 의미하고, 기준 자료를 계통적으로 추출한 영향이 큰 것으로 생각된다.

SPOT의 ISDd는 MSS나 TM에 비해 작고, ISDs는 1.0보다 크다. 이는 SPOT이 약간의 군집을 가진 임의분포임을 의미한다(그림 2 참조). 실제로 SPOT의 분류결과는 좌하단에 약간의 밀집이 보이며 오분류 화소의 분포도 기준 자료의 계통적인 형태와는 다소 거리가 있음을 알 수 있다(그림 5).

밀집 분포를 보인 SPOT영상에 대하여 오차 행렬(Confusion matrix) 분석[5]을 실시하였고(표 3) SPOT 분류 결과로부터 오분류 영역을 추출하였다.

표 2. 분석 결과

영상항목	MSS	TM	SPOT
전체 정확도	0.603	0.622	0.639
Kappa 계수	0.542	0.713	0.470
ISDd	0.745	0.753	0.702
ISDs	0.975	0.879	1.165

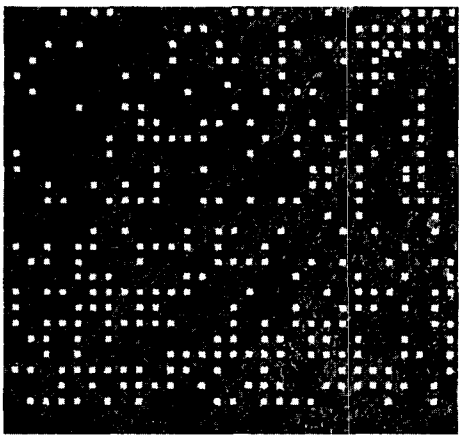


그림 3. MSS의 오분류 화소

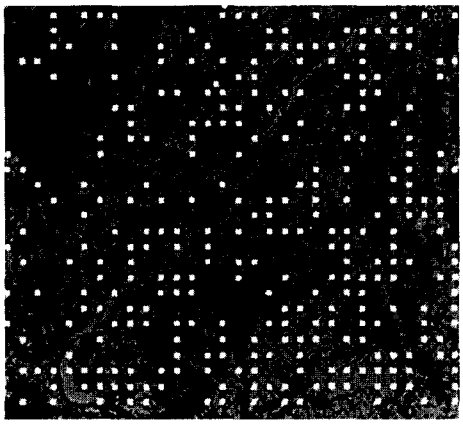


그림 4. TM의 오분류 화소

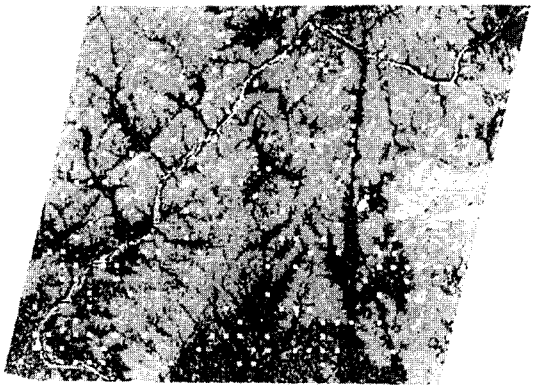


그림 5. SPOT의 오분류 화소

표 3. SPOT의 오차 행렬

분류결과 \ 기준데이터	1	2	3	4	5	6	오분류
1(논)	204	4				1	5
2(밭)	27	23	26		5	25	83
3(주거)		3	25		4	1	8
4(호수)				8		2	2
5(강)					24		0
6(산림)	31	43	25	2	4	764	105
오류	58	50	51	2	13	29	203

오분류 영역을 추출하기 위해 5,000 m×5,000 m 크기의 window가 사용되었고, 그에 따라 한번의 window에 25개에서 36개까지의 기준 데이터가 포함될 수 있었다.

오분류 영역은 기각역보다 낮은 전체 정확도를 가진 window를 찾아 추출하였다(그림 6). 본 연구에서는 기각역을 전체정확도 0.5로 하였다. 추출된 오분류 영역에

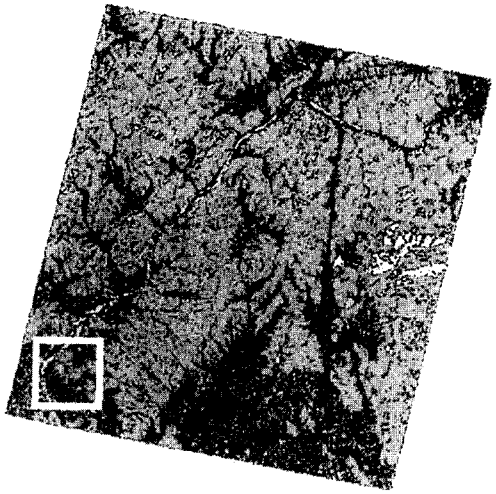


그림 6. 추출된 오분류 영역

표 4. 클래스 별 면적 구성 (면적:km^2)

Class	오분류 영역		전체 영상	
	면적	비율	면적	비율
논	12.707	0.508	315.056	0.310
밭	2.199	0.088	17.156	0.017
주거지	1.938	0.078	16.352	0.016
호수	0.070	0.003	1.589	0.002
강	0.350	0.014	5.171	0.005
산림	6.786	0.271	614.672	0.605
미분류	0.950	0.038	46.542	0.045
계	25.00	1.000	1016.538	1.000

대한 각 분류항목의 비율과 원 분류 결과의 전체 면적에 대한 각 분류항목별 면적의 비율을 비교하면 표 4와 같다.

추출된 오분류 영역에 대한 밭, 논 및 주거지의 비율이 원 분류결과에 비해 현저히 높음을 알 수 있다. 이는 밭, 논 및 주거지가 서로 혼동되고 있다는 의미를 가지며, 이는 오차행렬에서도 상당 수의 밭, 논 및 주거지가 기준 데이터와 다름을 통해서도 확인할 수 있다. 203개의 오분류 화소 중 밭에 해당하는 것이 83개로 절대적인 양에서는 산림에 비해 적은 수이나, 상대적으로 오분류된 화소가 옳게 분류된 화소에 비해 많은 수를 가지고 있다.

이는 영상 데이터 취득 시기와 토지피복의 규모에 의해 밭, 논 및 주거지의 분광특성이 흐려져 있기 때문이라고 판단된다. 영상이 취득된 시기는 4월 5일로 연구대

상 지역의 밭에는 아직 대부분의 밭작물의 파종이 이루어지지 않았고, 논에는 아직 물대기가 시작되지 않은 상태이다. 따라서, 밭과 논은 거의 나지 상태로 유사한 분광특성을 보이며, 특히 밭의 경우 봄 가뭄으로 함습도가 부족하여 주거지와 같은 분류항목과 분광특성에서 큰 차이를 보이지 않게 된다. 이와 같은 상태는 특히 밭, 논 및 주거지가 혼재된 지역에서 더 큰 영향을 미치게 된다.

실제로 본 연구에서 추출된 오분류 영역은 연구 대상 지역에서 밭, 논 및 주거지의 혼재가 가장 심한 지역으로 전체 정확도가 0.5보다 작았다. 따라서, 오분류영역은 기각되거나 재분류되어야 한다.

5. 결 론

본 연구를 통하여 얻은 결론은 다음과 같다.

첫째, 위성 영상 분류오차의 공간적 분포 특성을 평가하기 위해 ISDd와 ISDs를 제안하여, 분류오차의 공간분포 평가를 위한 지표를 얻을 수 있었다.

둘째, ISDd와 ISDs를 이용하여 복잡한 지형이나, 토지피복 또는 영상의 잡음에 의해 발생하는 오분류 화소의 국지적 밀집을 평가할 수 있게 되었다.

셋째, ISDd와 ISDs를 이용하여 sub-image 수준에서 분류 결과의 채택과 기각을 결정할 수 있게 되었다.

넷째, 공간적인 특성은 대상지역 전체에 대해 균일하지 않기 때문에, 추출된 오분류 영역의 공간적 특성을 상세히 조사하는 것에 의해, 토지 피복, 분류 항목의 설정, 분광 신호, 그림자의 영향 등, 오차행렬을 이용한 분류 오차의 발생 원인에 대한 분석의 신뢰성을 높일 수 있게 되었다.

향후 연구 과제로는 통계적 검정, 기각 영역의 자동 추출및 선형성을 가진 오차를 보이는 mixel 또는 잡음의 영향에 대한 평가가 있다.

참고문헌

1. 김용일, "인공위성영상의 분류정확도 향상에 관한 연구 -지형적 분광반사특성 보정과 통계적 분류기법을 중심으로," 서울대학교 박사학위논문, 1991.
2 박민호, "정준상관분석을 이용한 영상분류기법의 개발", 서울대학교 박사학위논문, 1996.
3. 이희연, "지리통계학," 법문사, 1991, p. 659.
4. Congalton, R. G. and Mead, R. A., "A Quantitative Method to Test for Consistency and Correctness in Photointerpretation", Photogrammetric Engineering & Remote Sensing, Vol. 49, No.1, 1983, pp. 69~74.
5. Lillesand, T. M. and Kiefer, R. W., "Remote Sensing and Image Interpretation", Wiley, 1994.
6. Rosenfeld, G. H. and Fitzpatrick-Lins, K., "A Coefficient of Agreement as a Measure of Thematic Classification Accuracy", Photogrammetric Engineering & Remote Sensing, Vol. 52, No. 2, 1986, McGraw Hill, pp. 223~227.
7. Spiegel, M. R., "Statistics 2/ed.", McGraw Hill, 1992, .
8. Upton, G. and Fingleton, B., "Spatial Data Analysis by Example", Vol. 1, Wiley 1985.