

수치위성자료를 이용한 광역의 토지피복분류 기법 Land Cover Classification Techniques for Large Area using Digital Satellite Data

박 병 옥*
Park, Byung-Uk

요 旨

Landsat TM 자료를 이용하여 전라남도 전역에 대한 토지피복분류를 실시하므로써 광역 분석시 자주 발생하게 되는 서로 다른 두 path 데이터의 통합 분석기법을 제시하였다. path가 다를 경우 관측일이 다르므로 개별적 분석 후 통합하였으며, 이 과정에서 두 scene 간의 분류가능 항목의 변화 및 자료 중복지역의 처리문제 등이 대두되었다. 또한 광역 분석시 위도차에 따른 데이터값의 변화로 인하여 분류항목 및 training field의 선정이 어려웠다. 본 연구에서는 수차례의 시행착오를 통하여 이러한 문제에 대한 해결방안을 제시하였으며, 분류 정확도 향상을 위한 Bayesian 최대우도법 및 majority 필터링의 효율성을 입증하였다.

ABSTRACT

This paper is to provide land cover classification techniques for large area ranged in different paths by classifying Landsat TM data of Jeonnam province. The analyses proceeded by individual scene because acquired dates are not same in different paths. In this processing, troubles had happened something like variation of classes can be classified in two scenes and choice problem about overlapped area. Since spatial effects in large area affect data values, it was difficult to make a selection of classes and training fields. we could present a solution about these problems by trial and error method, and found that Bayesian maximum likelihood classification and majority filtering were effective to improve classification accuracy.

1. 서 론

인공위성을 이용한 원격탐사 자료는 항공사진에 비하여 광역적이고 주기적인 자료획득이 가능하고 수치 형태로 취득할 수 있다는 장점으로 그 활용에 대한 연구가 활발히 이루어지고 있다. 우리나라에서도 그동안 도시지역의 토지피복분류, 산림자원 관리, 환경자원 관리 등을 중심으로 다수의 연구가 진행되어 그 효용성이 입증되었다.^{4,6)} 최초의 자원탐사위성인 Landsat의 발사 이래 SPOT, ERS 등 지구자원탐사를 위한 인공위성의 개발이 지속적으로 이루어지면서 인공위성 자료의 공간해상력이 향상되고 있으며, 관측기법의 다양화로 입체영상 및 레이다영상 등을 취득할 수 있게 됨에 따라

앞으로 그 활용은 더욱 실용적이고 다양화될 것이며,⁷⁾ 아울러 GIS의 주 자료원으로서의 역할이 크게 강조될 것으로 예측된다.

이러한 원격탐사 자료는 래스터 형태로 저장되는데, 래스터 자료는 벡터 자료에 비하여 자료구조가 단순하고 공간분석이 용이한 반면 자료의 양이 방대하다는 단점이 있다. 특히 인공위성 자료는 매우 방대하며 Landsat TM(Thematic Mapper) 자료의 경우 1 scene의 크기가 270MB나 된다. 따라서 전체 자료를 동시에 처리하기에는 컴퓨터의 처리용량 및 속도에 많은 제약을 받게 되고, 지금까지 대부분의 연구는 소규모 지역을 대상으로 이루어지게 되었으며, 이것이 원격탐사 활성화의 한 장애요인으로 작용하였다. 그러나 최근 컴퓨터 처리능력의 급속한 발달은 일반 PC 수준에서도 TM full-scene을 처리할 수 있게 되었으며, 이에 따라 대규

*안성산업대학교 토목공학과 조교수

모 지역을 대상으로한 연구가 가능하게 되었다.

따라서 본 연구에서는 PC를 이용하여 Landsat TM 자료로부터 전라남도 전지역에 대한 토지피복 정보를 추출하고자 하였다. 이 과정에서 과거 소규모 지역을 대상으로 할때에는 문제가 되지 않았던 많은 문제점들이 도출되었으며, 여러번의 시행착오를 거쳐야 했다. 본 연구에서는 전라남도 지역의 토지피복분류를 통하여 대규모 지역에 대한 수치위성자료를 분석하는 과정 및 고려사항을 제시하고자 한다.

2. 전처리 과정

2.1 사용 자료

전라남도 지역의 토지피복현황을 분석하기 위하여 Landsat-5호의 TM 자료를 사용하였다. TM 자료는 공간해상력이 30m이며(밴드 6은 120m), 총 7개 밴드로 구성되어 있고 각각의 밴드는 주제별 분석에 적합하도록 세분화되어 있다.

TM full-scene은 5965 line x 6792 column으로 구성되며 약 170 km(중) x 185 km(횡) 지역을 포함하는데, 연구대상 영역인 전라남도 지역은 그림 1과 같이 위성 궤도 115-36의 약 3/4 및 116-36의 약 1/2에 중복되어 있다. 따라서 최근의 자료중 관측일이 비슷한 두개의

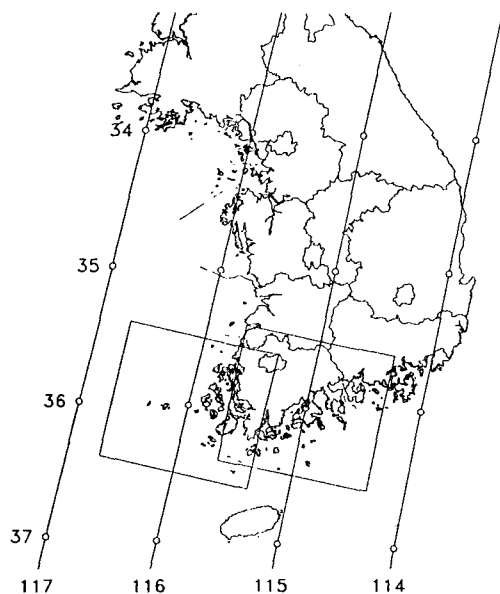


그림 1. Landsat-5호의 관측도

full-scene(115-36 : 1994년 6월 1일, 116-36 : 1994년 7월 26일)을 분석하였으며, 처리시스템으로는 PC-486(16MB RAM, 2.5GB HDD) 및 ERDAS 소프트웨어를 사용하였다.

2.2 GCP 선정

위성의 궤도는 원형의 극궤도를 이루지만 위성이 움직이는 동안 지구가 자전하기 때문에 실제궤도는 약간 경사지게 되며, 이로 인하여 관측된 위성자료는 실제 지형위치에 대해 왜곡된다. 따라서 GCP(Ground Control Point, 지상좌표를 확인할 수 있는 화상의 기준점)를 이용하여 정밀 기하보정을 실시한 후 분석하여야 한다.³⁾

GCP는 화상에서 명확히 식별되는 점이어야 하며 주로 도로교차점, 육지와 수면의 경계점 등이 사용되고, 또한 기복변위의 영향을 최소화하기 위하여 가급적 동일 표고의 점을 선정하는 것이 좋다. GCP의 분포는 전체적인 결과에 영향을 미치므로 해석대상 영역 전반에 고르게 분포되어야 하며, 가급적 많을수록 좋은 결과를 얻게 된다.

본 연구에서는 115-36 scene에서 26개, 116-36 scene에서 14개의 GCP를 선정하였으며, 보다 정확을 기하기 위하여 도로폭이 좁은 농로간의 교차점을 주로 선정하였다. 이중 9개의 GCP는 두 scene의 중복지역에 분포하는 공통 GCP로서, 이는 최종 결합시 두 scene의 매끄러운 결합을 위하여 동일지점을 사용하였으며, 전반적인 분포는 그림 2와 같다.

GCP에 대해서는 화상좌표 및 지상좌표를 각각 측정하고 이 결과를 이용하여 변환식을 결정하게 된다. 화상좌표는 모니터 상에서 마우스와 커서를 이용하여 라인 및 컬럼 좌표를 읽어 결정하였으며, 이때 소프트웨어의 확대 기능을 이용하여 1/8 화소단위의 정밀도로 관측하였다. 지상좌표는 TM 자료의 해상력이 30m인 점을 감안하여 1/25,000 지형도로 부터 독취하였다.

2.3 기하보정

GCP의 지상좌표를 (X,Y), 화상좌표를 (U,V)라 하면 변환식은 $X=f(U,V)$, $Y=f(U,V)$ 의 형태를 이루며 일반적으로는 식(2-1)과 같은 1차 affine 변환식이 사용된다. 그러나 affine 변환식은 축척변환 및 방향계수, 회전, 원점이동 등만을 표현하므로 지구의 곡면현상을 설명하

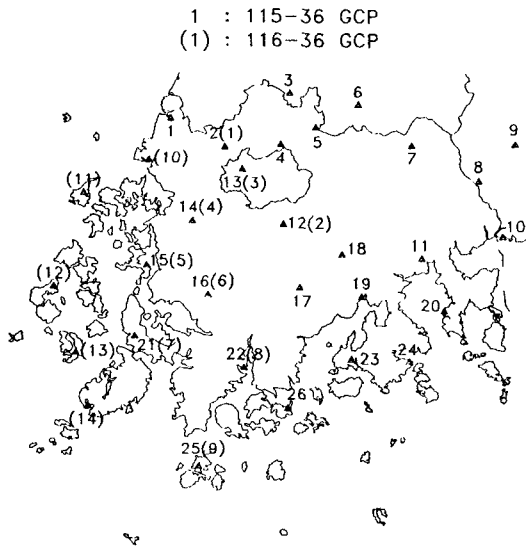


그림 2. GCP 분포도

표 1. 기하보정시 다항식 차수의 영향

다항식 차수	미지수 (최소 관측점수)	1 화소당 연산 회수	RMSE (115-36)	RMSE (116-36)
affine 변환식	6개 (3개)	8회	0.90	0.83
2차 다항식	12개 (6개)	26회	0.68	0.62
3차 다항식	20개 (10개)	58회	0.60	0.48
4차 다항식	30개 (15개)	108회	0.49	-
5차 다항식	42개 (21개)	180회	0.32	-

지 못한다. 따라서 본 연구에서와 같이 대상지역이 매우 넓어 곡면현상을 고려해야 할 경우에는 고차다항식을 사용해야 하며, 본 연구에서는 식(2-2)와 같은 2차 다항식을 적용하였으며 미지계수는 최소제곱법을 적용하여 결정하였다.

$$\begin{aligned} X &= a_0 + a_1U + a_2V \\ Y &= b_0 + b_1U + b_2V \end{aligned} \quad (2-1)$$

$$\begin{aligned} X &= a_0 + a_1U + a_2V + a_3UV + a_4U^2 + a_5V^2 \\ Y &= b_0 + b_1U + b_2V + b_3UV + b_4U^2 + b_5V^2 \end{aligned} \quad (2-2)$$

실제 선정된 GCP에 대하여 affine 변환식 및 2차 이상의 고차 다항식을 적용하였을 때 평균제곱근오차(RMSE)는 변환식의 차수가 높을수록 감소하여 정밀변환을 할 수 있음을 알 수 있다. (표 1 참조) 그러나 차수

가 높아질수록 미지계수의 수가 증가하고 이에 따라 기하보정에 소요되는 컴퓨터 계산시간이 급증하게 된다. 따라서 RMSE가 급격히 줄어드는 2차 다항식이 가장 효율적이며, 또한 측정한 지상좌표 및 화상좌표의 오차를 감안할때 2차 다항식의 적용이 타당함을 알 수 있다. 2차 다항식을 적용하였을때의 RMSE는 각각 0.68, 0.62 화소로서 GCP의 선정작업이 매우 양호하게 이루어졌다.

기하보정 후의 화상은 지상좌표계를 기준으로 일정한 거리 간격으로 재배열 된다. TM 자료의 해상력은 30m이지만 지상 수신소에서 사용자에게 공급하는 자료는 일련의 처리과정을 거쳐 사방 약 28.5m를 나타내게 되므로 기하보정시의 화소간격은 28.5m가 되도록 하였다.

재배열 과정은 출력화소의 위치를 식(2-2)의 역변환식에 대입하여 미보정 화상에서의 화상좌표를 구한 후 그 위치에 해당하는 화소값을 취하므로써 이루어진다. 이때 계산된 화상좌표는 화소단위 이하이므로 주위의 화소값을 참조하여 최종 데이터값을 결정하게 되는데, 여기에는 최근린 내삽법(nearest neighbour), 공일차 내삽법(bilinear interpolation), cubic convolution 방법 등이 적용된다.

최근린 내삽법은 계산된 화상좌표에 가장 가까운 화소의 값을 택하는 것으로 가장 빠르지만 출력화상이 거칠게 된다. 공일차 내삽법은 인접한 4개 화소값에 대하여 거리에 따른 가중평균치를 택하는 것으로, 비교적 부드러운 화상을 얻을 수 있으며 계산시간도 그리 많이 소요되지는 않아 일반적으로 가장 많이 사용되는 방법으로서 본 연구에서도 이 방법을 적용하였다. cubic convolution 방법은 주위 16개 화소에 대하여 sinh 함수 등을 적용하는 방법으로서 가장 선명한 화상을 얻을 수 있지만 계산시간이 많이 소요된다.

2.4 해석대상 영역의 절출

기하보정된 화상에 대하여 연구대상 영역인 전라남도 지역의 자료만을 절출하였다. 이를 위하여 전라남도의 행정경계를 1/50,000 지형도에서 디지털화 하였다. 전라남도는 내륙지역 뿐만 아니라 바다와 무수히 많은 섬으로 이루어져 있는데, 바다의 경계가 명확하지 않고 또한 본 연구가 육지지역의 토지피복현황 파악이 주 목적이므로 내륙지역 및 주요 섬(약 200여개) 지역

만을 절출하였다. 한편 내륙지역은 하나의 거대 폴리곤으로서 매우 많은 점으로 디지털화되어 한번에 처리할 수 없었으며, 이에 따라 여러개의 폴리곤으로 분할하여 절출한 후 접합시켰다.

본 연구와 같이 해석대상 영역이 서로 다른 path에 걸쳐 있는 경우 인접 path간에 상당한 중복지역이 발생하는데, 이 경우 중복부분에 대한 데이터 취사선택의 문제가 발생한다. 색조 모자이크의 경우에는 양쪽 scene의 색조를 비슷하게 만든 후 지성선을 중심으로 각각 절출하는 것이 타당하나, 토지피복분류의 경우 두 scene의 관측일 차이에 따른 피복상황의 변동이 존재하므로 대상영역의 주된 분포 scene을 최대한 반영하는 것이 타당할 것이다. 그러나 한반도 전역의 분석 등과 같이 매우 많은 scene을 사용할 경우에는 지성선에 의한 접합이 바람직하다.⁹⁾

따라서 본 연구에서는 중복지역에 대해 115-36 scene을 최대한 활용하고 부족 부분만을 116-36 scene에서 절출하였으며, 이때 섬 지역이나 반도 부분에서 경계가 이루어질 경우에는 그 전체를 116-36 자료로 사용하여 경계부분에서의 토지피복상황의 변동에 따른 이질감을 최소화 하고자 하였다. 한편 TM scene의 좌우측 양단부분의 데이터는 밴드에 따라 자료의 시작 및 끝부분의 차이가 존재함을 관측할 수 있었는데, 밴드 1,2,3,4,7,5,6의 순으로 오른쪽으로 이동되어 있으며 전체적인 이동 크기는 약 200 켈럼 정도였다. 따라서 토지피복분류시 scene의 좌우측 양단의 약 200 켈럼 부분의 데이터는 활용할 수 없다는 것을 고려하여야 한다.

양쪽 scene에서 행정경계를 따라 절출해 본 결과 전라남도 좌측상단 부분의 데이터가 약간 부족하여 115-35 및 116-35 scene의 일부가 추가로 필요함을 발견하였으나 데이터 확보의 어려움 때문에 간과하기로 하였다.

3. 토지피복분류

토지피복분류는 일반적으로 사전정보(ground truth data)를 이용하는 감독분류(supervised classification)를 적용하며, 이것은 사용자가 각 분류항목의 특성을 지정하는 training 단계와 이를 이용한 분석단계로 이루어진다.

training 단계에서는 화상에서 각 분류항목별로 확실한 영역(training field)을 지도나 현장조사 등을 통하여 설정하고, 그 지역의 데이터 분포를 기준으로 하여 분

류항목별 통계특성 자료를 작성하게 된다. training field는 이후 분류단계의 결과에 직접적으로 영향을 미치므로 신중히 선택해야 하며 보통 시행착오 방법을 거치게 된다.¹⁰⁾

분석단계에서는 위의 통계특성 자료를 이용하여 일정한 기준에 의하여 모든 화소들을 순차적으로 분류하게 된다. 이때 분류 기준설정 방법에 따라 최단거리법(minimum distance classifier), 평행육면체법(parallelepiped classifier), 최대우도법(maximum likelihood classifier) 등이 적용된다.

본 연구에서는 컴퓨터 처리시간이 많이 걸리지만 가장 정확한 분류방법인 최대우도법을 적용하였다. 이 방법은 training field의 데이터가 정규분포를 따른다는 가정하에 각 분류항목의 평균벡터와 공분산행렬을 계산한 후, 확률밀도함수를 이용하여 각각의 화소를 확률이 가장 높은 분류항목으로 분류하는 방법이다. 분류작업은 각 scene 별로 수행하였으며, 분류후 하나로 통합하였다.

3.1 분류항목 및 Training Field의 선정

일반적으로 Landsat TM 자료로 분류할 수 있는 항목은 USGS의 토지피복분류체계로 볼때 level II까지 가능하다.¹¹⁾ 그러나 이때 그 종류는 자료의 취득시기 및 대상지역에 따라 다르게 나타난다.

본 연구에 사용한 자료는 시기적으로 볼때 약 2개월 정도 차이가 나지만 환경의 변화속도가 활발한 여름에 관측되어 두 scene에서 분류할 수 있는 항목의 종류가 상당히 다르게 나타났고, 따라서 분류 후의 결과의 통합을 고려하여 최종적인 분류항목은 level I로 결정하였으며 산림, 논, 밭, 도시, 물, 나지, 비닐하우스, 염전 등 총 8개 항목으로 설정하였다. 이 중 비닐하우스 항목은 주로 논이나 밭 지역에 나타나고 계절의 영향에 상당히 민감하게 변화하므로, 본 연구에서처럼 통합분석을 하는 경우에는 바람직하지 않으나 데이터값의 특성이 너무 독특하게 나타나므로 하나의 항목으로 설정할 수 밖에 없었다. 그리고 염전은 서해안쪽에만 존재하여 116-36 scene의 분석에서만 적용되었다.

한편 해석영역이 매우 넓은 관계로 level I으로 분류작업을 실시하면 특정항목의 분산이 너무 커져 다른 항목에까지 영향을 미치는 현상이 나타났다. 즉 최대우도법으로 분석할 경우 분산이 높은 항목일수록 그 항목의

표 2. Training Field의 평균값 및 (표준편차)

분류항목	화 소 수	band 1	band 2	band 3	band 4	band 5	band 6	band 7
활엽수	746	68.26(2.79)	30.13(1.50)	24.65(1.13)	153.37(17.40)	102.52(9.55)	153.58(3.85)	26.66(2.54)
침엽수	729	69.19(3.29)	29.30(2.34)	26.72(3.23)	73.56(8.31)	64.78(16.72)	161.17(9.25)	21.49(6.95)
잡초	143	74.03(2.72)	38.28(2.19)	31.49(1.80)	159.78(17.05)	108.75(11.13)	157.15(2.11)	32.59(3.86)
논	398	82.73(4.22)	40.18(3.36)	46.96(6.22)	50.43(6.63)	26.18(7.49)	164.57(4.00)	10.16(2.68)
밭	141	96.38(4.89)	50.46(3.67)	67.88(6.61)	91.58(6.32)	152.60(10.38)	185.70(4.22)	80.33(7.98)
도시	648	118.02(9.84)	58.78(6.71)	78.10(10.94)	74.07(9.07)	122.83(16.21)	189.94(5.88)	75.42(10.08)
깊은물	1002	79.00(6.77)	32.00(5.78)	26.35(6.49)	11.38(1.20)	6.53(0.99)	145.68(2.48)	3.00(0.81)
얕은물	379	99.15(13.43)	52.09(8.13)	63.53(14.01)	28.33(5.44)	10.07(4.81)	149.00(4.08)	4.56(2.57)
오염된물	192	74.14(3.30)	31.30(3.22)	29.47(4.04)	20.58(9.85)	15.24(7.91)	154.44(2.00)	6.45(2.95)
일반나지	578	155.35(11.06)	89.95(8.06)	134.74(20.96)	124.72(19.44)	216.99(31.42)	180.66(8.62)	128.98(16.36)
간척나지	524	129.40(3.85)	67.27(2.44)	91.03(3.76)	77.37(3.41)	150.32(6.55)	170.39(3.83)	93.91(5.03)
산지나지	180	80.93(4.87)	37.78(2.94)	48.02(5.23)	73.83(12.47)	130.13(13.61)	193.51(4.91)	61.67(6.47)
습한나지	414	90.50(5.65)	40.69(2.96)	47.75(4.68)	40.16(3.09)	53.51(5.43)	169.03(5.45)	23.19(4.16)
비닐하우스	149	119.51(11.82)	58.32(5.67)	71.84(7.67)	100.26(8.00)	106.38(9.53)	168.34(2.18)	45.70(5.54)

로 분류될 확률이 상대적으로 높아져 예기치 못한 결과를 초래하며 분류항목간 분산값의 적절한 균형이 요구되었다. 특히 산림의 경우 위도에 따라 많은 차이를 보였고 도서지역과 내륙지역간에도 많은 차이가 나타나 단일항목으로 분류할 경우 분산이 매우 높게 나타남을 알 수 있었다. 논, 밭의 경우도 모내기 시기 또는 추수시기에 관측한 자료라면 위도에 따라 매우 다양하게 나타날 것이다. 때로는 분류항목의 분산이 적절하더라도 계절적인 영향으로 분류항목 간의 평균값이 비슷하게 나타나 이의 구별을 위해 항목을 세분해야 되는 경우도 발생하였다.

따라서 실제 분류작업시에는 각 scene 별로 특성에 맞게 level II 또는 그 이상의 수준까지 세분화하여 분류하였으며, 그 항목 수의 결정에는 training field에 따른 예비분석 등 수차례의 시행착오가 요구되었다. 115-36 scene에서는 산림 및 물을 3가지로, 나지는 4가지로 세분하였으며, 116-36 scene에서는 산림을 4가지, 논을 2가지, 나지를 3가지로 세분하였다. 115-36 scene을 예로 들면, 산림의 경우 대략 활엽수, 침엽수, 잡초 등으로 구분하였고, 물의 경우에는 당시 논, 밭의 상태가 물로 채워져 있어 물과 논, 밭의 구분을 위하여 깊은 물, 얕은 물, 도심지의 오염된 물 등으로 세분화 하였으며, 나지의 경우에는 운동장과 같은 나지, 간척사업 후 마른 나지, 산지의 나지, 습한 나지 등으로 나누었다.

각 분류항목 별로 그 항목을 가장 잘 대표할 수 있는 training field를 선정하였으며, 이 과정은 1/25,000 지형도 및 칼라합성시의 색조 등을 참고로 하여 이루어졌고 원칙적으로 대상영역에 고루 분포하도록 하였다. training field의 타당성은 일차적으로 자체에 대한 confusion matrix를 작성해 봄으로써 판단할 수 있으며, 전체 분석을 통하여 최종적인 판단을 하게 되는데 바람직한 결과가 나올때까지 검토 및 수정을 반복하게 된다. 해석 영역이 넓은 관계로 본 연구에서 사용한 시스템으로 한번 분석하는데는 약 12시간이 소요되었으므로 작업의 효율을 기하기 위하여 500 x 500 정도의 소규모 지역을 3-4개 선정하여 예비분석을 실시하면서 training field를 수정하였다. 116-36 scene에서는 분석 대상이 대부분 서해안 섬 지역이지만 분석의 균형을 이루기 위하여 중북부분에 위치한 내륙지역에서도 training field를 선정하였다. 115-36 scene의 training field에 대한 밴드별 평균 및 표준편차값의 분포는 표 2와 같다.

3.2 토지피복분류

표 2에서 보면 활엽수와 침엽수는 밴드 4, 5에서 가장 잘 구분되며, 일반나지는 거의 모든 밴드에서 다른 분류항목들보다 평균값 및 표준편차가 높게 나타나고, 물의 경우 가시광선대(밴드 1~3)에서 물의 특성에 따

른 차이가 발생하는 반면 적외선 파장대(밴드 4~7)에서는 다른 항목보다 매우 낮게 나타나 확실히 구분될 수 있음을 알 수 있다. 그리고 열적외선 밴드인 밴드 6의 경우 대체적으로 평균값이 비슷하지만 어떤 분류 항목 간에는 많은 차이가 나타나 분류작업시 도움이 될 수 있음을 알 수 있다.

또한 표 2에서 활엽수와 잡초, 깊은 물과 오염된 물 등이 상대적으로 비슷한 분포를 나타내고 있으나, 최대우도법은 단순히 평균값과 표준편차만으로 분류하는 것이 아니라 각 밴드간의 공분산행렬을 이용하므로 평균값이 비슷하더라도 분류항목간 구분이 가능하다.

본 연구에서는 토지피복분류를 위하여 사전확률을 이용하는 Bayesian 최대우도법을 적용하였으며 7개 밴드 모두 사용하여 분류하였다. Bayesian 최대우도법은 임의의 화소 X가 분류항목 ω_i 에 속할 확률 $P(\omega_i | X)$ 이 가장 높은 항목으로 분류하는 것이다.²⁾

$$P(\omega_i | X) = P(X | \omega_i) / P(X) \quad (3-1)$$

여기서 $P(X) = \sum_{i=1}^m P(X | \omega_i) P(\omega_i)$

$P(X | \omega_i)$: 분류항목 ω_i 에서 X가 나타날 확률

$P(\omega_i)$: 분류항목 ω_i 의 사전확률

식(3-1)에서 $P(X)$ 는 개별 분류항목과 관계 없으므로, $P(X | \omega_i) P(\omega_i)$ 만으로 판단할 수 있으며 이것을 판별함수로 한다. 한편, $P(X | \omega_i)$ 는 정규확률밀도함수로 표현하면 n차원에 대하여 식(3-2)로 표현된다.

$$P(X | \omega_i) = \frac{1}{(2\pi)^{n/2} |Cov_i|^{1/2}} \exp[-0.5(X-M_i)^T(Cov_i)(X-M_i)] \quad (3-2)$$

여기서 M_i : 분류항목 ω_i 의 평균벡터

Cov_i : 분류항목 ω_i 의 공분산행렬

$|Cov_i|$: 공분산행렬의 행렬식

판별함수에서는 상수값을 가감승제하거나 대수를 취하더라도 관계 없으므로, 식(3-2)에 대하여 자연대수를 취하고 상수항을 소거하여 간단히 하면 판별함수 $D_i(X)$ 는 식(3-3)으로 표현된다.^{8,12)}

$$D_i(X) = \ln P(\omega_i) - 0.5 \ln(|Cov_i|) - 0.5 (X-M_i)^T(Cov_i)(X-M_i) \quad (3-3)$$

즉, 임의의 화소 X에 대하여 각 분류항목별로 식(3-3)에 의한 값을 구한 후 이 중 값이 가장 높은 항목으로 분류하게 된다. 여기서 사전확률 $P(\omega_i)$ 은 분류항목간의 가중치로 생각할 수 있으며, 자체적으로 추정하기가 어렵기 때문에 확실하지 않은 경우 보통 모두 1을 적용한다. 그러나 임의 항목으로 분류하기가 애매한 화소의 경우 이 값에 따라 최종적인 분류항목이 달라질 수 있으므로 일차적인 분류결과에 따라 적절한 가중치를 적용함으로써 분류의 정확도를 향상시킬 수 있다.

본 연구에서도 training field의 조정으로도 미진한 부분을 항목간 가중치를 다르게 부여함으로써 해결하고자 하였으며, 적절한 가중치를 찾기 위하여 반복 분류작업을 실시하였다. 각 scene별로 Bayesian 최대우도법으로 분류한 후 이것을 앞서 언급한 level I의 8개 분류항목으로 재변환(recode) 하므로써 두 scene 간의 분류항목을 일치시켰다. 다음으로 두 scene의 분류결과를 통합하므로써 전라남도 전체의 토지피복분류 화일을 작성하였다.

한편 Landsat-5호는 사용 자료의 관측당시 발사된지 10여년이 경과한 상태여서 scene에 잡음이 비교적 많았고, 또한 임의 집단내에 단일 화소로 표현된 이질적인 토지피복상황은 TM 자료의 해상력에 비추어 불매 혼합화소(mixel) 현상에 의한 오분류일 가능성이 높기 때문에 이러한 요인을 소거하기 위하여 필터링(filtering)을 최종적으로 실시하였다. 분류결과 화일의 화소값들은 분류항목 순으로 일련 번호의 값을 가지므로, 필터링시 평균값이나 중앙값으로 대체하는 방법은 예기치 못한 결과를 유발하게 된다. 따라서 윈도우내

0	1	0
1	1	1
0	1	0

그림 3. 반경 1의 원형 윈도우

가장 다수의 값으로 대체하는 majority 방법을 적용하였으며, 이때 윈도우의 크기 및 형태로는 반경 1의 원형 윈도우를 채택하였다. (그림 3 참조)

3.3 분류결과에의 해석

분류결과에의 신뢰도는 일차적으로 사용된 training field에 대한 confusion matrix에 의해 평가할 수 있다. 그 결과 115-36 scene에서 99.6%(표 3 참조), 116-36 scene에서 99.3%로 나타나 training field의 선정 및 분류항목 간 분리가 양호함을 알 수 있다.

전라남도에 대한 토지피복분류의 최종적인 결과는 그림 4 및 표 4와 같다. 그림 4에서 산림은 녹색, 논은 노란색, 밭은 갈색, 도시는 적색, 물은 청색, 나지는 흑색, 비닐하우스는 담청색, 염전은 분홍색으로 표현하였다.

분류결과에의 비교를 위하여 1994년 현재 전라남도 및 광주광역시 지역의 통계연보 자료를 종합 정리하여 표 5에 나타냈다.¹⁾⁵ 통계연보의 자료는 지적공부상 지목별 현황이기 때문에 토지피복을 대상으로한 본 연구의 분류 체계와 다르므로 적절한 항목끼리 묶었다. 한편, 본 연구에서 전라남도 좌측상단 부분 및 작은 섬지역의 일부

누락에도 불구하고 연구결과에의 총 면적이 통계연보의 면적보다 크게 나타난 것은, 통계연보에서는 도서지역에 대해 종합면적만 제시하고 지목별 현황이 누락되어 있어 도표에 포함시킬 수 없었기 때문이며 그 면적을 더하면 전라남도(광주 포함)의 총 면적은 13,574.53 km²이다.

표 4와 5의 항목별 퍼센트 비율을 비교해 볼때 논이 통계자료 보다 훨씬 적게 나타난 것은 비닐하우스의 대부분이 논에서 발생했기 때문으로서 경작형태의 변화를 알 수 있었으며, 비닐하우스 비율과 합하면 비슷한 값을 나타낸다. 그리고 지적공부상 농촌지역에 위치한 소규모 대지 및 개간된 임야 등이 토지피복분류 시 밭으로 분류되어 밭은 상대적으로 높게, 산림과 도시는 상대적으로 약간 낮게 나타났다. 물의 경우 해남군 및 영암군 일대의 간석작업 후 발생한 대규모 나지 때문에 분석결과 물의 비율이 줄어들고 나지가 많이 분류되었다. 이상의 결과를 종합해 볼때 만족할 만한 분류 정확도를 얻었다고 판단되며, 인공위성 자료를 이용한 토지피복분류가 변화하는 토지의 실질적 이용 현황을 파악하는데 통계자료보다 효율적임을 알 수 있었다.

표 3. 115-36 Scene에서의 Training Field에 대한 Confusion Matrix

분류항목	f1*	f2	f3	pd	df	ur	w1	w2	w3	b1	b2	b3	b4	vn
화 소 수	746	729	143	398	141	648	1002	379	192	578	524	180	414	149
분류 결과	f1	746	0	0	0	0	0	0	0	0	0	0	0	0
	f2	0	729	0	0	0	0	0	0	0	0	0	0	0
	f3	0	0	143	0	0	0	0	0	0	0	0	0	0
	pd	0	0	0	398	0	0	0	0	2	0	0	0	1
	df	0	0	0	0	141	4	0	0	0	0	0	1	0
	ur	0	0	0	0	0	641	0	0	0	7	2	0	0
	w1	0	0	0	0	0	0	1001	0	0	0	0	0	0
	w2	0	0	0	0	0	0	1	373	0	0	0	0	0
	w3	0	0	0	0	0	0	0	6	190	0	0	0	0
	b1	0	0	0	0	0	0	0	0	0	571	0	0	0
	b2	0	0	0	0	0	0	0	0	0	0	522	0	0
b3	0	0	0	0	0	0	0	0	0	0	0	179	0	
b4	0	0	0	0	0	0	0	0	0	0	0	0	413	
vn	0	0	0	0	0	0	0	0	0	0	0	0	0	149
정확도(%)	100	100	100	100	100	98.9	99.9	98.4	99.0	98.8	99.6	99.4	99.8	100

* f1: 활엽수, f2: 침엽수, f3: 잡초, pd: 논, df: 밭, ur: 도시, w1: 깊은물, w2: 얕은물, w3: 오염된물
b1: 일반나지, b2: 간척지나지, b3: 산지나지, b4: 습한나지, vn: 비닐하우스



그림 4. 전라남도 토지피복분류 결과 (1994년)

표 4. 전라남도 토지피복분류 결과 (1994년)

분류 항목	화소수	면적(km ²)	퍼센트(%)
산림	8,792,904	7,142.04	56.1
논	1,696,807	1,378.23	10.8
밭	2,313,529	1,879.16	14.8
도시	598,129	485.83	3.8
물	424,821	345.06	2.7
나지	642,850	522.15	4.1
비닐하우스	1,152,795	936.36	7.4
염전	39,763	32.30	0.3
합 계	15,661,598	12,721.13	100.0

표 5. 전라남도 토지지목현황 (통계연보, 1994년)

분류 항목	면적(km ²)	퍼센트(%)
산림(임야, 공원)	7,363.76	59.6
논	2,156.71	17.4
밭 (전, 과수원, 목장용지, 묘지)	1,411.71	11.4
도시(대지, 공장용지, 도로, 철도용지)	607.59	4.9
물 (하천, 구거, 유지)	614.29	5.0
나지(잡종지, 학교용지, 체육용지, 제방, 기타)	152.16	1.2
염전	58.42	0.5
합 계	12,364.64	100.0

4. 결 론

본 연구에서는 Landsat TM 자료를 이용하여 전라남도 전체에 대한 토지피복분류를 실시하므로서 대규모 지역에 대한 분석시 발생할 수 있는 문제점을 고찰해

보고, 이의 해결 방안을 모색하였으며 다음과 같은 결론을 얻었다.

1. 두 path에 걸친 지역에 대한 토지피복분류시 분석 절차 및 고려요소, 접근방안 등을 제시하므로서 수치위

성자료로 광역 분석시 발생할 수 있는 최대 문제점에 대한 해결 방안을 제시할 수 있었다.

2. 대규모 지역에 대한 기하보정식으로는 2차 다항식의 적용이 가장 바람직하다.

3. 광역의 분석시에는 위도차에 따른 데이터 값의 변화가 많이 발생하므로 분류항목의 선정 및 training field의 선정에 각별한 주의가 요망된다.

4. 사전확률을 이용하는 Bayesian 최대우도법을 적용 하므로써 분류 정확도를 향상시킬 수 있었다.

5. 분류결과에 대하여 반경 1의 원형 윈도우를 사용한 majority 필터링을 적용하므로써 데이터의 잡음 및 혼합화소의 영향을 소거할 수 있었다.

감사의 글

본 연구에 필요한 TM 자료 및 ERDAS 소프트웨어를 제공하여 주신 환경부 및 캐드랜드(주)의 관계자 여러분께 감사드립니다.

參考文獻

1. 광주광역시, 제35회 광주통계연보, 1995, pp. 62-63.
2. 안철호, 김용일, "인공위성 원격탐사 데이터의 분석 정확도 향상에 관한 연구", 한국측지학회지, 제9권 제2호, 1991, pp. 81-91.
3. 안철호, 박병욱 "LANDSAT-5의 TM과 MSS 데이터의 비교에 관한 연구", 한국측지학회지, 제4권 제2호, 1986, pp. 27-41.
4. 안철호, 윤정섭, 안기원, "도시 및 지역연구를 위한 인공 위성 Data의 분석 및 그 활용방안에 관한 연구", 한국측지학회지, 제3권 제1호, 1985, pp. 1-13.
5. 전라남도, 제35회 전남통계연보, 1995, pp. 74-77.
6. 최승필, 홍성태, 박재훈, "Landsat TM 자료를 이용한 임종구분에 관한 연구", 한국측지학회지, 제11권 제1호, 1993, pp. 55-60.
7. Brisco, B. and Brown, R.J., "Multidate SAR/TM Synergism for Crop Classification in Western Canada", PE & RS, Vol. 61, No. 8, 1995, pp. 1009-1014.
8. Brown, N. and Smith C., ERDAS Field Guide, ERDAS Inc., 1990, pp.156-157.
9. Fuller, R.M., Groom, G.B. and Jones A.R., "The Land Cover Map of Great Britain : An Automated Classification of Landsat Thematic Mapper Data", PE & RS, Vol. 60, No. 5, 1994, pp. 553-562.
10. Hixon, M., Scholz, D. and Fuhs, N., "Evaluation of Several Schemes for Classification of Remotely Sensed Data", PE & RS, Vol. 46, No. 12, 1980, pp. 1547-1553.
11. Lillesand, T.M. and Kiefer, R.W., Remote Sensing and Image Interpretation, John Wiley & Sons, 1994, pp. 169-178.
12. Swan, P.H. and Davis S.M., Remote Sensing : The Quantitative Approach, McGraw-Hill, 1978, pp. 148-158.