

원격탐사 영상의 분류정확도 향상을 위한 인공지능형 시스템의 적용

An Application of Artificial Intelligence System for Accuracy Improvement in Classification of Remotely Sensed Images

양인태* · 한성만** · 박재국 ***
Yang, In Tae · Han, Soung Man · Park, Jae Kook

요 旨

이 연구는 원격탐사 영상의 분류정확도를 향상시키기 위한 방법으로써 신경망 이론과 퍼지집합이론을 각각 적용하였다. 원격탐사 영상은 토지피복도, 식생도, 지질도 등 주제도를 만드는데 많이 이용되고 있다. 원격탐사 영상의 감독분류에 대한 정확도는 트레이닝 지역의 선정, 분류항목의 할당 문제로 인해 많은 차이를 보인다. 일반적인 영상 분류법은 영상 내의 모든 영상소가 균질하다고 가정한다. 그러나, 이러한 가정은 영상내의 수많은 혼합 영상소를 분류해내는 데에는 적합하지 않다. 이러한 문제를 극복하기 위해 퍼지 집합이론을 적용하였으며, 퍼지 집합이론의 멤버십을 이용하였다. 퍼지 집합이론은 하나의 영상소를 멤버십의 정도에 따라 여러 가지 항목으로 분류할 수 있는 장점이 있다. 그러나, 퍼지분류법과 통계학적인 분류법은 화소값의 분포가 비정규적일 때 좋지 않은 분류 결과를 나타내며 처리 시간이 늦고 많은 컴퓨팅 비용이 드는 단점이 있다. 그 대안적인 방법으로서 신경망분류법을 들 수 있는데, 신경망 분류법은 비모수적 분류법으로서 일반적인 분류기법보다 좀 더 좋은 결과를 나타내고 있고, 한번 트레이닝 되면 빠르게 데이터를 분류할 수 있다.

ABSTRACT

This study applied each Neural Networks theory and Fuzzy Set theory to improve accuracy in remotely sensed images. Remotely sensed data have been used to map land cover. The accuracy is dependent on a range of factors related to the data set and methods used. Thus, the accuracy of maps derived from conventional supervised image classification techniques is a function of factors related to the training, allocation, and testing stages of the classification. Conventional image classification techniques assume that all the pixels within the image are pure. That is, that they represent an area of homogeneous cover of a single land-cover class. But, this assumption is often untenable with pixels of mixed land-cover composition abundant in an image. Mixed pixels are a major problem in land-cover mapping applications. For each pixel, the strengths of class membership derived in the classification may be related to its land-cover composition. Fuzzy classification techniques are the concept of a pixel having a degree of membership to all classes is fundamental to fuzzy-sets-based techniques. A major problem with the fuzzy-sets and probabilistic methods is that they are slow and computational demanding. For analyzing large data sets and rapid processing, alternative techniques are required. One particularly attractive approach is the use of artificial neural networks. These are non-parametric techniques which have been shown to generally be capable of classifying data as or more accurately than conventional classifiers. An artificial neural networks, once trained, may classify data extremely rapidly as the classification process may be reduced to the solution of a large number of extremely simple calculations which may be performed in parallel.

1. 서 론

원격탐사 분야에서 토지피복분류는 가장 대표적이고

전형적인 응용 방법의 하나로, 산림, 물, 농경지 등과 같은 지표면의 물리적 상황을 여러 가지 응용기법을 통하여 분류하는 것이다. 일반적으로 많이 이용되는 분류법으로 최대우도분류법이 있는데, 이것은 한 영상소가 어떤 항목에 포함될 가능성에 근거한 방법이다. 이 분류방법은 Bayes의 결정규칙을 따르며 모든 분류항목의 분포가 정규분포라는 가정 하에 개발된 방법이므로 비정규

*강원대학교 토목공학과 교수

**강원대학교 토목공학과 박사수료

***강원대학교 토목공학과 박사과정

분포일 때는 좋지 못한 분류 결과를 가져온다. 그리고, 처리 시간이 길고, 입력 밴드의 수가 늘어날수록 계산 시간도 증가하는 단점을 가지고 있다.³⁾

이러한 최대우도분류법의 단점을 보완하기 위한 대안으로 신경망 이론을 제안할 수 있다. 신경망은 비모수적, 비선형적 특성을 가지며, 간단한 기능을 갖는 단수의 처리기들이 인접한 처리기들과 정보를 주고받으면서 동작하는 형태로 이루어지기 때문에 병렬 처리를 할 수 있다는 특징도 가지고 있다. 이러한 특성은 토지피복분류에 있어서 연산을 분산 병렬 처리할 수 있게 함으로써 처리 시간을 단축시킬 수 있다. 그러나 신경망은 학습 기능과 대량 병렬성을 보유한 유연성이 뛰어난 기법이지만, 한 영상소내에서의 각 항목들에 대한 분류를 실행할 수 없다는 단점이 있다.¹⁾

이런 단점으로부터 유용한 정보를 추출해 낼 수 있는 것이 Zadeh교수에 의해 제안된 퍼지 집합이론(Fuzzy Set Theory)이다. 신경망 이론이 위에서 설명한 비모수적, 비선형적이라는 장점을 갖는 반면에, 영상 분류에 있어서의 이분법적인 할당을 배제시킬 수가 없다. 그러나, 퍼지집합 이론은 불명확한 원격탐사 영상의 영상소값들에 할당 가능한 분류항목들을 모두 포함시킴으로써 최대우도분류법이나 신경망분류법에 비해 다양한 분류 항목들을 분류해 낼 것이다. 이 연구에서는 분류법에 따른 결과를 실제 측정된 참조 자료와 상호 비교하여 그 정확도를 검증하였다.

2. 연구내용 및 방법

2.1 연구지역의 선정 및 방법

이 연구에서는 토지피복군의 분포가 다양하면서도 트레이닝 및 검증에 위한 표본추출이 비교적 용이함으로써 분류기법의 연구에 좋은 조건을 가지고 있는 지역을 선정하기 위해 1996년 영상획득 당시부터 지금까지 도시 개발에 따른 변화가 거의 없는 춘천시 소재의 일부 지역을 선정하였다. 또한 이 연구에서 이용한 연구의 흐름도는 그림 2와 같다.

2.2 전처리

2.2.1 최적사용 밴드의 선정

이 연구에서는 통계학적 사상 선택 방법 중 모든 밴드 조합에 대하여 사용 밴드 사이의 상관 정도를 도식적으

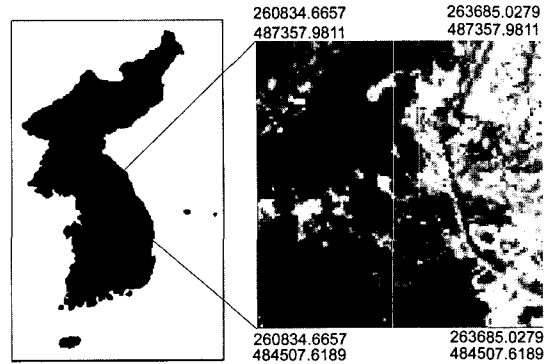


그림 1. 연구지역

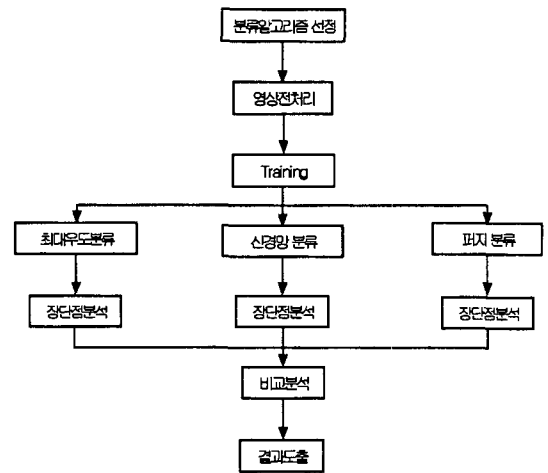


그림 2. 연구의 흐름도

표 1. 분기도 분석 결과

밴드 조합	평균 분기도	분류항목간 조합				
		1	1	1	2	2
234	2000	2000	1999	1999	2000	2000
123	1907	1962	1999	1936	1999	1186

※분류항목 : 1=삼림, 2=물, 3=인공구조물, 4=농경목초지

로 측정함으로써 가장 상관 정도가 정량적으로 큰 밴드 조합을 찾아낼 수 있게 해주는 사상공간도표 방법을 이용하여 토지 피복 분류에 가장 적합한 밴드를 선정하였다. 표 1은 Landsat TM 1, 2, 3, 4, 5, 7 밴드를 선택하여 분류항목간 분기도를 측정된 결과로서 연구 대상지역을 주로 구성하고 있는 육역과 수역을 잘 구분하고, 식생에 대해서 판별능력이 우수한 2, 3, 4 밴드를 선정

하였다.^{5),7)}

2.2.2 기하보정(Geometric Correction)

미보정 영상을 좌표변환시키기 위해서는 다항방정식이 이용된다. 영상의 기하 왜곡 정도에 따라 지상기준점의 수를 결정하고 다항방정식의 차수를 결정한다. 보통 위성 영상의 좌표변환에는 1차 또는 2차 방정식이 많이 이용되며, 이 연구에서는 1차원 Affine 변환방정식을 이용하여 변환을 실시하였다. 환경부 홈페이지에서 제공한 지상기준점 8지점을 선정한 후 원격탐사 영상에 대해 Map to Image 방식으로 기하보정을 하였다. 원격탐사 영상의 RMSE는 9.38 m로 Landsat TM 영상의 일반적인 허용 오차 15 m를 만족시켰다. 다음은 영상의 Affine 변환행렬의 계수이다.

$$\begin{pmatrix} 27.99 & -5.26 & 96203 \\ -5.28 & -28.01 & 541926 \end{pmatrix} \quad (2-1)$$

1차 방정식은 다음과 같이 표현할 수 있다.

$$\begin{aligned} X &= 27.99x - 5.26y + 96203 \\ Y &= -5.28x - 28.01 + 541926 \end{aligned} \quad (2-2)$$

재배열은 영상소값의 손실이 가장 적은 최근린 내삽법으로 하였으며, 한 영상소의 크기는 28.5 m×28.5 m로 결정하였다. 재배열이 완료된 후 100×100의 연구지역을 절취하였다.

2.2.3 트레이닝 표본 선정 및 통계 계산치 추출

전처리 이후에 연구지역으로 선정된 영상 내에서 어떠한 분류 항목을 선정하여 분류할 것인가를 결정하였다. 우선 눈으로 식별이 가능한 삼림, 물, 인공구조물, 농경목초지, 나대지 순으로 분류항목을 결정하였다.

나대지의 경우 인공구조물과 섞여있어 구분이 잘 되지 않았으나 항공사진과 실제 현장 검증을 통하여 트레이닝 표본으로 선정할 수 있었다. 인공구조물은 주거지역, 상업지역, 도로, 교량 등 사람이 인위적으로 만든 대상을 인공구조물로 선정하였으며, 농경목초지는 시기적으로 농경지와 목초지의 구분이 잘 되지 않을 뿐만 아니라 연구지역 내에 목초지라고 할 만한 지역이 거의 없어 이를 따로 분리하는 것은 분류 결과에 좋지 않은 영향을 미칠 것으로 사료되어 이 연구에서는 이를 통합하여 선정하였다.

트레이닝 기법으로는 서로 이질적인 분광특성을 가진 영상소들을 주어진 매개변수에 따라 자동적으로 표본추

표 2. 분류항목의 트레이닝 통계값

분류항목	영상	band2	band3	band4
	삼림	min	19	17
max		25	30	83
mode		20	19	65
Std. Dev		1.26	2.62	8.16
Frequence		100		
물	min	24	22	9
	max	39	43	19
	mode	25	25	13
	Std. Dev	4.09	5.13	2.20
	Frequence	188		
인공구조물	min	31	39	44
	max	51	77	77
	mode	38	59	62
	Std. Dev	4.13	7.79	6.69
	Frequence	109		
농경목초지	min	24	23	94
	max	34	43	134
	mode	28	27	122
	Std. Dev	2.27	4.38	9.30
	Frequence	62		
나대지	min	48	77	69
	max	105	169	144
	mode	64	112	92
	Std. Dev	12.64	20.98	18.40
	Frequence	51		

출을 해주는 Region Growing 기법을 이용하여 표본을 선정하였다. 연구지역이 작은 관계로 인해 너무 많은 샘플링이 되지 않도록 영상소값의 범위를 작게하여 추출하였다.

표 2는 트레이닝 결과로 나타난 분류항목의 평균벡터와 표준편차 등의 통계 계산치이다.

3. 토지피복분류

3.1 최대우도 분류법에 의한 토지피복분류

토지피복분류를 실시하기 위해서 연구 지역에서 수집이 용이한 대표적인 트레이닝 표본(5가지 : 삼림, 수역, 농경지, 인공구조물, 나대지)을 Region Growing 기법을 이용하여 트레이닝 표본으로 선정하였다. 또한 최대우도

표 3. 최대우도분류법의 분산-공분산 행렬

분류항목	밴드	band2	band3	band4
	삼림	band2	1.596	2.659
band3		2.659	6.870	6.046
band4		4.293	6.046	66.507
물	band2	16.726	20.479	5.562
	band3	20.479	26.327	7.239
	band4	5.562	7.239	4.858
인공구조물	band2	17.086	31.219	9.676
	band3	31.219	60.606	16.774
	band4	9.676	16.774	44.769
농경목초지	band2	5.152	8.535	-8.254
	band3	8.535	19.146	-27.267
	band4	-8.254	-27.267	86.467
나대지	band2	159.738	256.256	191.416
	band3	256.256	439.973	352.356
	band4	191.416	352.356	338.661

분류 결정규칙에 이용하기 위해 분류 각 항목간의 공분산 행렬, 평균벡터, 표준편차 등의 통계 계산치를 추출하였다. 그리고 최대우도 분류기법을 이용해서 얻어진 분류항목별 분류면적과 분류결과 영상을 출력하였으며 분류정확도 검증용 실시하였다.

이 연구에서는 먼저 트레이닝 작업이 완료된 후 각각의 영상소를 최대우도 결정규칙 알고리즘에 따라 분류를 실시하기 위해 표 3에 나타난 분류항목의 공분산 행렬을 이용하여 전개값과 역행렬을 계산한 후 전체 영상소를 대상으로 개개 분류항목에 할당할 확률을 계산하였으며, 확률이 가장 큰 항목으로 해당 영상소를 할당하여 분류를 실시하였다.

표 3은 분류과정 중에서 얻어진 각 분류항목의 공분산 행렬이다.

최대우도 분류법에 의한 토지피복분류 결과는 삼림이 1.493 km², 물이 1.473 km², 인공구조물이 2.875 km², 농경목초지가 1.496 km², 나대지가 0.785 km²의 면적분포로 나타났다.

이 연구의 목적이 분류 결과의 정확도를 검증하는 것이므로 실제 현장 검증과 항공사진을 통해 항목별 면적을 산출하는 것이 바람직하지만 연구 여건상 너무 많은 시간과 비용이 소요됨으로 전체 연구지역에 대한 면적 산출보다는 여러 가지 분류항목들로 구성된 지역을 선정

하여 가로, 세로를 5등분하여 검증을 실시하였다. 검증방법은 항공사진을 이용하여 면적을 추출하였고 실제 현장 검증을 통하여 항공사진에 나타난 지표 사상들의 분포를 확인하였다.

3.2 신경망 이론에 의한 토지피복분류

신경망 이론을 이용한 인공위성 영상의 분류는 영상소 값의 분포가 정규분포를 따르지 않더라도 일반적인 최대우도 분류보다 좀 더 정확한 분류 결과를 나타낸다. 학습동안에 신경망은 사상공간 내에서 축소된 경계구역의 형태를 갖추게 된다. 이러한 성능은 영상소값의 분포에 상관없이 경계구역이 학습 알고리즘에 의해 반복적으로 조정되기 때문에 통계적인 방법보다 좋은 결과를 나타낸다. 신경망 이론의 감독 분류의 적용은 일반적인 분류기법과 유사한 점이 많은데, 차이점이 있다면 학습과 분류 알고리즘에 미묘한 차이가 있다. 신경망의 학습 단계는 최대우도분류의 항목간 평균과 공분산 행렬의 계산과 유사하다. 그러나 통계적인 기법들은 계산이 한 번에 끝나는 대신에, 신경망은 반복적으로 학습을 하게된다. 이 연구에서 사용한 백프로퍼게이션 알고리즘은 목표한 최소 오차가 원하는 결과값과 신경망의 실제 결과값 사이에서 얻어질 때까지 반복된다.¹⁾²⁾

신경망을 학습시킨다는 것은 몇 가지 초기 파라메타들을 결정하는 것을 포함하는데, 이 연구에서 사용된 신경망 구조에 있어서 처음에 결정해야 하는 파라메타는 입력 밴드와 분류 항목, 입력과 출력의 오차값, 학습률, 히든 층의 수이다. 앞에서 열거한 파라메타를 바꿔 줌으로써 사용자가 원하는 결과치에 근접한 영상을 얻을 수 있다. 여기에서 사용자가 원하는 영상이란 목표치에 수렴해야 하며 시간이 오래 걸리지 않아야 한다.⁶⁾

이 연구에서 수행한 연구방법을 설명하면 다음과 같다. 첫 번째 단계에서 트레이닝 데이터를 결정해야 하며, 트레이닝 데이터의 원하는 항목을 선정해야 한다. 트레이닝 데이터 선정 방법은 일반적인 최대우도와 같은 방법으로 수행되며, 행렬로 표현되기 때문에 위치에 대한 정보도 필요하게 된다. 이 연구에서는 Matlab으로 프로그램을 만들어서 데이터를 처리하였는데, Matlab에서는 데이터를 행렬로 처리할 수 있도록 하기 때문에 원격탐사 영상을 처리하기 위해서는 행렬로의 변환과정이 필요하였다. 그런다음 전체적인 신경망의 구조를 정의해야 한다. 학습이 시작되었을 때, 신경망의 모든 가

중치들은 임의의 값을 갖게 된다. 그렇지 않으면 신경망은 최소의 학습오차에 수렴하지 않을 수도 있다.¹⁹⁾ 그런 다음 학습률이 설정되어야 한다. 초기의 학습률이 트레이닝의 성공에 중요한 영향을 미치는 것은 아니지만, 학습률이 너무 높게 설정되면 네트워크가 불안정해지기 때문에 트레이닝의 속도가 감소한다. 실제로 프로그램 실행 중에 같은 오차를 가진 경우에 학습률이 높게 되면 수렴이 되지 않거나 학습 반복 횟수가 늘어나는 경우를 확인할 수 있었다. 마지막으로 요구되는 파라미터는 실험적으로 결정되어야 하는 학습 수렴 경계치이다. 원하는 값과 실제 결과값 사이에 최소 제곱 오차와 같은 경계값이 설정되어야 한다. 이런 경계값이 만족되었을 때 학습이 완료된 것이다. 신경망이 학습 데이터에 대해서 훈련이 잘 되었다면 영상의 나머지 부분까지 작용하지 않았을 것이고, 반대로 학습이 충분히 이루어지지 않았다면 만족할 정도로 항목들을 분류해 낼 수 없을 것이다.⁶⁾

3.3 퍼지집합 이론에 의한 토지피복분류

원격탐사 영상의 한 영상소값은 그 지형에 대한 분광 특성을 나타내며 일반적으로 한 영상소는 순수한 영상소(pure pixel)라기 보다는 여러 가지 혼합된 영상소(Mixed pixel)로 나타난다.⁶⁾ 일반적으로 하나의 영상소가 하나의 영역에만 속한다면 그 분광특성과 일치하는 하나의 영역으로 분류되며, 이 영상소는 그 영역에 대한 Membership Grade Value가 1이 되고, 다른 영역에 대한 Membership Grade Value가 0이 된다. 따라서 분류하고자하는 영역을 퍼지 집합으로 나타내므로써 하나의 영상소는 하나의 영역에 어느 정도 속하고 다른 영역에도 어느 정도 속하는지 표현할 수 있다. 문제는 각 영상소가 어떤 영역에 어느 정도 속해 있는지를 결정하는 것이다. 이 과정을 Membership Modification이라고 하는데 다음의 순서대로 Membership Grade Value를 계산하고 퍼지 분류 영상을 생성할 수 있다.

이 연구에서 사용한 연구지역은 하나의 영상소내에 여

표 4. 항목 분리도의 분석

분류항목	영상	영상		
		band2	band3	band4
삼림	min	19	17	42
	max	25	30	83
	분리도	물:24~25 농경목초지:24~25	물:22~30 농경목초지:23~30	인공구조물:44~77 나대지:69~83
물	min	24	22	9
	max	39	43	19
	분리도	삼림:24~25 농경목초지:24~34	인공구조물:39~43 농경목초지:23~43	
인공구조물	min	31	39	44
	max	51	77	77
	분리도	물:31~39 농경목초지:31~34 나대지:48~51	물:39~43 농경목초지:39~43 나대지:77	삼림:44~77 나대지:69~77
농경목초지	min	24	23	94
	max	34	43	134
	분리도	삼림:24~25 물:24~34 인공구조물:31~34	물:23~43 인공구조물:39~43	나대지:94~134
나대지	min	48	77	69
	max	105	169	144
	분리도	인공구조물:48~51	인공구조물:77	인공구조물:69~77 농경목초지:94~134

리 가지 분류 항목을 가지고 있으며 분류 항목간 영상소 값의 분포가 중첩되어 있는 모습을 볼 수 있다. 최대우도 분류기법이나 신경망 이론을 이용한 분류기법의 경우에 이러한 영상소내의 이질성에 대한 관계를 설정하기 어렵기 때문에 하나의 항목으로 해당 영상소를 결정하는 오류를 범하게 된다. 그 결과 분류 결과의 정확도를 신뢰할 수 없게 된다.

분리도 분석 결과는 사용 밴드별로 분류항목간의 분광 특성이 유사한 정도를 보여주며, 퍼지 평균과 퍼지 공분산을 계산하는데 기초적인 자료가 된다.

밴드 1은 분리도가 좋은 반면에 밴드 2와 3은 중첩된 영상소값의 범위가 너무 커서 단일 밴드로 1차원 처리를 하였을 경우 많은 오차를 포함할 확률이 높다. 표 4의 사용 밴드별 분리도 분석에서처럼 개개 분류항목들은 일정 부분 인접한 분류항목과 영상소값들이 겹쳐져 있는 모습을 보여준다.

이 연구에서 미지의 영상소는 3차원 사상공간내의 위치에 비추어 봤을 때 여러 분류 항목에 걸쳐 일정 부분씩 겹쳐져 있다. 이러한 정보는 유용하게 사용될 수 있지만 일반적인 분류법에서는 무조건 하나의 분류항목으로 미지의 영상소를 할당함으로써 분류의 신뢰성에 영향을 미친다. 그러나, 퍼지 분류에서는 미지 영상소를 무조건 하나의 항목으로 할당하는 것이 아니라 Membership Grade를 할당하여 하나의 영상소를 세분화하여 분류할 수 있게 한다.

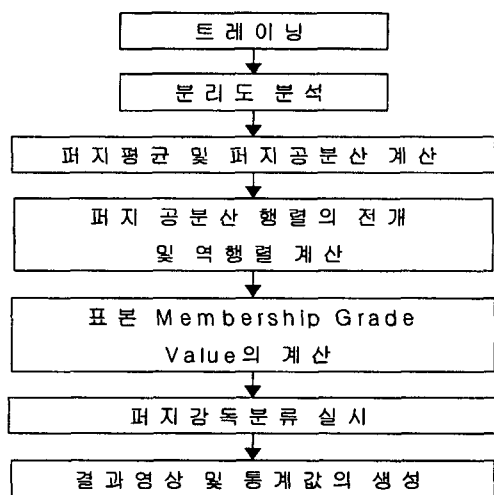


그림 3. 퍼지분류법의 순서도

3.3.1 표본 Membership Grade Value의 계산

퍼지 집합이론을 이용한 감독분류의 처음 단계는 표본들에 대한 Membership Grade Value를 계산하는 것이다. 먼저, 어느 분류항목이 인접 분류항목과 어느 정도 중복되는가를 계산한다. 3개 항목이 중복되는 경우에 Membership Grade Value를 계산하는 방식을 수식으로 표현하면 식 (3-1)과 같다.

$$\begin{aligned}
 \text{factor}_1 &= \text{bv_max}[j - 2] - k + 1 \quad (j = \text{분류항목 } k = \text{BV}) \\
 \text{factor}_2 &= k - \text{bv_min}[j - 1] + 1 \\
 \text{factor}_3 &= k - \text{bv_min}[j] + 1 \\
 \text{factor} &= \text{factor}_3 / (\text{factor}_1 + \text{factor}_2 + \text{factor}_3) \\
 \text{표본 Membership Grade Value} &= \text{factor} \quad (3-1)
 \end{aligned}$$

위 식에서

factor_1 : 첫 번째 분류항목의 중복가중치

factor_2 : 두 번째 분류항목의 중복가중치

factor_3 : 세 번째 분류항목의 중복가중치

bv_max : 최대 영상소값

bv_min : 최소 영상소값이다.

3.3.2 퍼지 평균과 퍼지 공분산 행렬의 계산

이 연구에서 분류방법의 하나로 이용된 최대우도분류 기법은 표본들에 대한 분류항목별 평균값과 공분산 행렬을 이용한다. 그러나 퍼지 함수를 이용한 분류기법에서는 일반적인 평균, 공분산 대신에 퍼지 평균과 퍼지 공분산 행렬을 계산하여야 한다.

기본적인 계산 기법은 동일하지만 이전 과정에서 계산한 표본 Membership Grade Value를 이용한다. 계산 공식은 다음과 같다.

퍼지평균

$$\mu_c^* = \frac{\sum_{i=1}^n f_c(x_i)x_i}{\sum_{i=1}^n f_c(x_i)} \quad (3-2)$$

퍼지 공분산 행렬

$$V_c^* = \frac{\sum_{i=1}^n f_c(x_i)(x_i - \mu_c^*)(x_i - \mu_c^*)^T}{\sum_{i=1}^n f_c(x_i)} \quad (3-3)$$

여기서, n은 전체 표본 영상소의 수

다음은 이 연구를 위해 작성한 프로그램에 의해 만들어진 퍼지 평균과 퍼지 공분산 행렬이다. 공분산 행렬은 사용영상의 차원의 수와 항목수에 따라 3×3행렬이 만들어진다. 최대우도분류법에서 사용되는 평균값과는 달리 퍼지분류법에서는 표본 Membership Grade Value를 이용하여 퍼지평균을 계산한다. 또한 퍼지 공분산은 Membership Grade Value에 대한 분류항목별 영상소 분포에 대한 분산 정도를 정량적으로 계산함으로써 미지 영상소가 각각의 분류항목에 속할 확률을 보다 객관적으로 분석할 수 있다.

여기서 계산된 퍼지 평균과 퍼지 공분산을 이용해 통계학적인 분류기법인 최대우도분류법에 적용하기 위해서는 먼저 각 분류항목의 영상소분포가 정규분포인가를 검증해야 한다. 영상에 대한 표본 분포와 중앙값, 중앙값에 대한 퍼지 평균의 근접도를 나타내므로써 분류항목별 정

표 5. 퍼지평균

분류항목	영상	band2	band3	band4
	삼림		20.9157	19.9245
물		27.3947	27.2826	12.5957
인공구조물		41.2973	58.8005	62.6767
농경목초지		29.0826	31.4224	114.3145
나대지		76.2271	126.0297	109.2698

표 6. 퍼지 분산-공분산 행렬

분류항목	영상	band2	band3	band4
	삼림	b2	1.1178	0.9267
b3		0.9267	3.0301	3.0764
b4		2.1350	3.0764	70.9168
물	b2	7.9523	7.9523	1.3494
	b3	2.5909	2.5909	1.5237
	b4	1.3494	1.3494	4.8579
인공구조물	b2	11.5067	18.7822	2.6956
	b3	18.7822	51.4492	5.2329
	b4	2.6956	5.2329	35.5464
농경목초지	b2	5.5222	3.8504	-3.5300
	b3	3.8504	21.0009	-12.9996
	b4	-3.5300	-12.9996	106.0507
나대지	b2	149.7851	235.0412	152.9881
	b3	235.0412	420.7615	284.0537
	b4	152.9881	284.0537	329.8160

규분포를 확인하였다.

3.3.3 Membership Grade Value의 생성

공분산 행렬의 전개와 역행렬의 생성이 완료된 후 확률값의 계산은 일반적인 최대우도 분류 알고리즘에 기초하여 수행하였다.

$$P_i^* = \frac{\exp[-0.5(x - \mu_i^*)^T V_i^{*-1} (x - \mu_i^*)]}{(2\pi)^{N/2} |V_i^*|^{0.5}} \quad (3-4)$$

N은 영상소 벡터의 차원이므로 이 연구에서는 사용 밴드의 수, 즉 3이다.

이러한 알고리즘은 다차원 원격탐사 영상의 영상소값 분포를 고려하여 보다 정밀한 기하학적 분리도를 구현하기 위해 일반적으로 이용되는 방법으로 분류에 사용된 원격탐사 영상의 영상소값이 정규분포를 가진다는 가정하에 만들어진 것이다. 그러므로 위의 알고리즘은 3차원 원격탐사 영상이 정규분포를 가진다는 전제하에만 유효하다. 이 연구에 이용된 Landsat TM 밴드 2,3,4 영상은 정규분포에 가까운 영상소분포를 가지고 있으므로 이러한 조건을 만족한다.

이전에 계산이 완료된 공분산 행렬의 전개값과 역행렬을 이 알고리즘에 적용하여 사용된 다차원 영상의 영상소 좌표별로 확률값을 계산하였다. 일반적인 최대우도 분류 알고리즘은 이 공식을 사용할 때 확률값이 가장 가까운 분류 항목에 영상소를 할당하게 된다. 하지만 퍼지분류법은 분류항목별로 구해진 각각의 확률값을 모두 더한 값으로 해당 분류항목의 확률값을 나누어 최종적인 Membership Grade Value를 생성한다. 다시 말하면, 확률값을 이용한 재배열 과정을 거침으로써 개개 분류항목에 할당될 Membership Grade Value를 계산하는 것이다. 그 알고리즘을 수식으로 단순하게 표현한 것이 식 (3-5)이다.

$$f_c(x) = \frac{P_c^*(x)}{\sum_{i=1}^m P_i^*(x)} \quad (3-5)$$

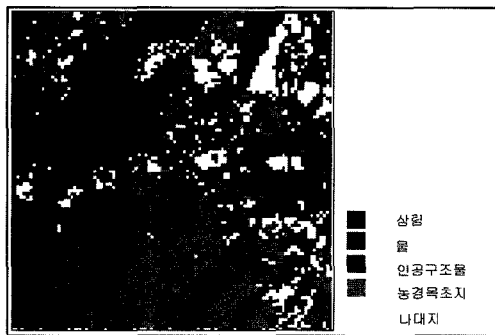
이렇게 분류항목 수만큼의 Membership Grade Value를 생성하였다.

퍼지집합 이론에 의한 토지피복분류 결과 삼림이 1.209 km², 물이 1.501 km², 인공구조물이 2.444 km², 농경목초지가 1.794 km², 나대지가 1.174 km²의 면적분포로 나타났다.

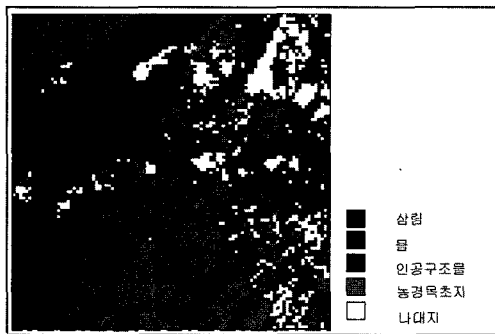
4. 검증 및 분석

4.1 결과 영상 및 검증 지점

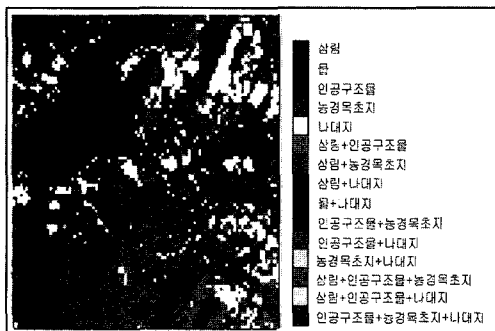
이 연구에서는 분류 결과의 정확도 검증을 위해서 항공사진을 이용하였다. 위성 영상은 지상기준점을 사용해서 기하보정을 하였으나 항공사진은 촬영 지역이 작고 해상도가 높기 때문에 1:1,000 수치지도를 이용하였다. 원격탐사 영상 및 항공사진은 10m 이하의 오차를 확인



<최대우도분류법>



<신경망분류법>



<퍼지분류법>

그림 4. 분류결과 영상

하였다. 그림 4는 각각의 분류 방법에 대한 전체 연구지역의 분류지도이다.

이 연구의 목표가 각각의 분류 기법에 대한 분류 정확도를 평가하는 것이 목적이므로 실제 검증 지점에 대한 항공사진을 가지고 검증을 실시하였다. 원격탐사 영상은 환경부 지상기준점으로 기하보정을 하였으며, 항공사진은 1/1,000 춘천시 수치지도를 이용해 기하보정을 실시하였다.

검증 지점으로 선정된 지역은 두 영상의 촬영시기가 15개월의 시간차가 있음에도 불구하고 변화가 거의 없는 지점으로 선정하였으며, 지상 사상의 판별이 용이한 곳으로 선정하였다. 그림 5는 연구지역 내에서 두 검증지점의 상태를 알기 쉽게 하기 위해 나타낸 것이다.

4.2 검증지점 1의 분류결과 및 분석

검증지점 1의 경우 삼림 오른쪽으로 그림자가 있는 것을 볼 수 있다. 그림자로 인하여 삼림과 물의 경계가 뚜렷하지 않아 각각의 면적을 산출하는데 있어서 오차가 생기는 것을 배제할 수 없으나 실제 현장 검증과 지형도를 통해서 경계를 설정하여 면적을 산출하였다.

그림 6은 검증지점 1의 항공사진을 원격탐사 영상의 크기와 동일하게 5×5로 나눈 모습이다.

그림 7은 검증지점 1의 분류결과를 각각의 분류법에 따라서 표현한 것이다.

표 7은 검증지점 1의 면적을 항공사진, 최대우도분류법, 신경망분류법, 퍼지분류법의 순서대로 항목별 면적을 표현한 것이다.

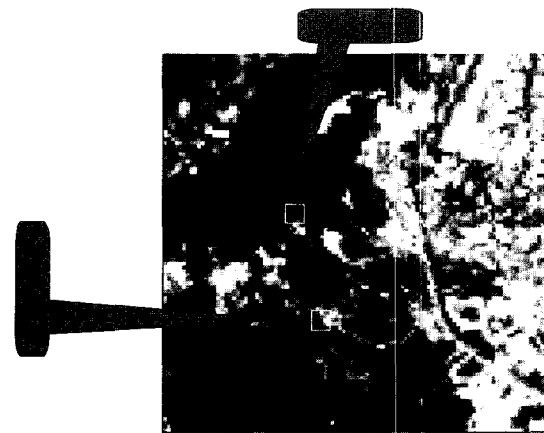


그림 5. 검증지점

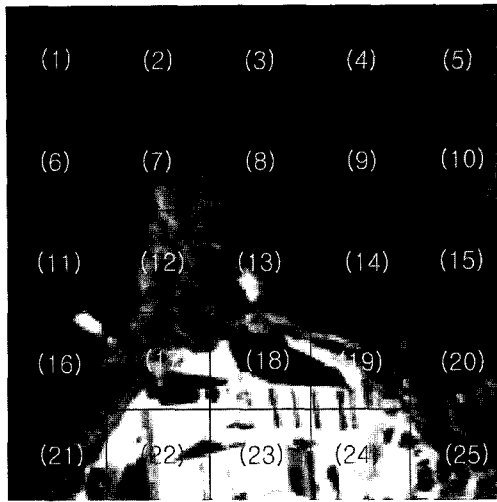


그림 6. 검증지점1의 항공사진

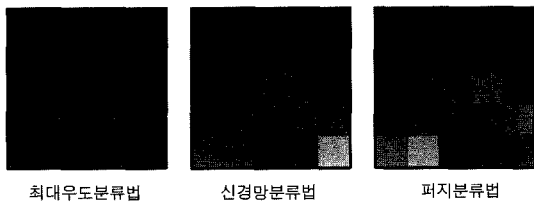


그림 7. 검증지점1의 분류결과

표 7. 검증지점1의 면적비교(단위:m²)

분류항목 분류기법	분류항목				
	삼림	물	인공 구조물	농경 목초지	나대지
항공사진	4,857	11,717	3,215	306	211
최대우도	4,061	10,559	5,686		
신경망	7,310	10,559	1,625	812	
퍼지	5,756	10,559	2,437	718	837

다음 설명은 검증지점 1에서의 영상소별 분류 결과가 다르게 나온 영상소를 중심으로 그 이유에 대한 설명을 정리하여 놓은 것이다.

검증지점 1에서 (7)번 영상소의 경우 항공사진을 보면 대부분의 지역이 삼림과 물로 이루어져 있으나 세 가지 분류방법이 모두 물로 표현되어 있다. 최대우도분류법과 신경망분류법에서는 차지하는 면적이 조금 더 큰 물로 할당되어 있어 물로 분류되었고, 퍼지분류법에서도 마찬가지로 물로 분류했으나 삼림의 MGV가 나타나지

않았다.

(14)번 영상소의 경우, 항공사진을 조사해 보면 물과 삼림의 분포를 볼 수 있으나 최대우도분류법의 분류 결과는 인공구조물로 분류되어 있고, 퍼지분류법에 의한 분류 결과 삼림과 나대지의 혼합지역으로 분류되어 있다. 이 영상소에 대한 퍼지분류법의 MGV는 삼림이 0.4089, 인공구조물이 0.0001, 나대지가 0.5910으로 추출되어 있어 나대지가 가장 많은 부분을 차지하고 있음을 알 수 있다. 그러나 항공사진의 조사 결과 나대지에 대한 존재는 발견되지 않고 있어 실제와는 다른 양상을 보이고 있다. 반면에 신경망분류법에 의한 분류 결과는 이 영상소를 삼림으로 처리하고 있어 실제 검증자료와 잘 맞는다는 것을 알 수 있다. 그러나, 퍼지에서는 삼림이 59% 이상인 것으로 되어 있는데 분기도 분석결과 삼림과 나대지의 분기도가 1729의 값을 갖는 것으로 보아 분기도가 좋지 않아 이러한 결과를 가져온 것으로 사료된다. 전체적인 삼림과 나대지의 분기도를 분석해보면 대부분의 밴드 조합에서 2000에 가까운 값을 갖는 것이 확인되었다. 위와 같은 이유로 인해 퍼지에 있어서 (14)번 영상소의 분류결과는 대표성을 띠다고 볼 수 없다.

(21)번 영상소의 경우, 항공사진에는 물, 삼림, 인공구조물들이 존재하고 있으며, 이 항목 가운데서 삼림이 가장 많은 부분을 차지하고 있음을 알 수 있다. 최대우도분류법에서는 이 영상소를 인공구조물로, 신경망분류법은 삼림으로 각각 처리하고 있어 최대우도분류법의 분류 결과보다는 신경망분류법의 분류결과가 잘 맞는다는 것을 확인할 수 있다. 반면에, 퍼지분류법에서는 이 영상소 값을 삼림과 나대지의 혼합지역으로 분류하고 있다. 퍼지분류법의 MGV에 대한 추출결과 삼림이 0.6777, 인공구조물이 0.0015, 나대지가 0.3208로 나타나고 있다. 이 지역에 대한 현장조사를 실시해 보니, 삼림과 인도블럭 등이 존재하고 있어 비교적 잘 맞는다고 할 수 있으나 나대지의 MGV가 인공구조물의 MGV보다 큰 값이 계산되어 다소 맞지 않는 결과를 얻었다.

(25)번 영상소의 경우, 항공사진과 현장조사를 통해서 알아본 결과 삼림, 잔디밭, 인도블럭으로 구성되어 있었으며, 이 가운데서 삼림과 잔디밭이 대부분을 차지하고 있고 그 사이사이에 인도가 나 있었다. 최대우도분류법에서는 이 지역의 대부분이 삼림과 잔디밭임에도 불구하고 인공구조물로 처리되었으며, 신경망분류법에서는 농경목초지에 해당하는 값으로 분류하고 있어 비교적 만족

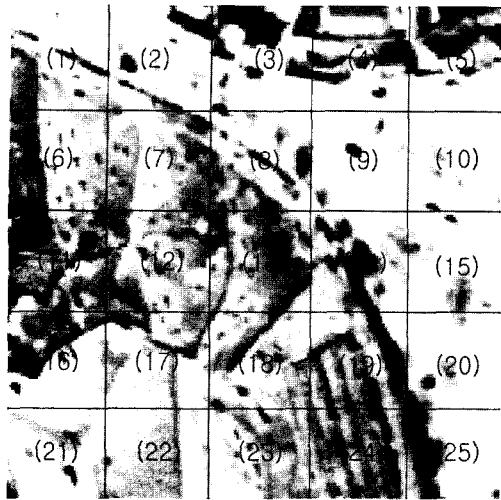


그림 8. 검증지점 2의 항공사진

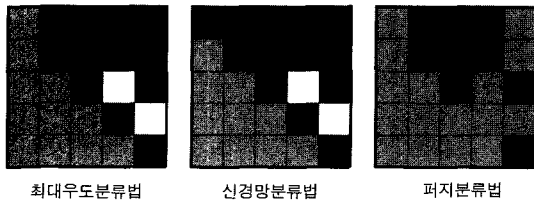


그림 9. 검증지점 2의 분류결과

한 값을 얻었다. 퍼지분류법의 경우, 실제로 삼림이 상당히 많은 부분을 차지하고 있음에도 불구하고 인공구조물과 농경목초지의 혼합지역으로 분류되었다. 퍼지분류법의 MGV를 살펴보면, 인공구조물이 0.9995, 나대지가 0.0005로 인공구조물이 대부분을 차지하고 있음을 나타내고 있어 항공사진과는 대조를 이루고 있다.

4.3 검증지점 2의 분류결과 및 분석

검증지점 2의 경우 항공사진에서 거의 수직으로 촬영된 영상이며 각 항목들 간의 경계가 명확하여 면적을 산출하기 쉬웠으며, 왼쪽지역은 가을 추수가 거의 끝난 상태라 나대지의 모습을 띠고 있다. 그림 8은 검증지점 2의 항공사진을 원격탐사 영상과 동일하게 5×5로 나눈 모습이며, 그림 7은 검증지점 2의 분류결과를 분류법에 따라서 나타낸 것이다.

표 8은 검증지점 2의 면적을 항공사진, 최대우도 분류법, 신경망분류법, 퍼지분류법에 대한 각 항목별 면적을

표 8. 검증지점 2의 면적비교(단위:m²)

분류항목	분류항목				
	삼림	불	인공 구조물	농경 목초지	나대지
항공사진	0	0	7,550	11,988	768
최대우도	0	0	9,747	8,935	1,624
신경망	0	0	10,559	8,123	1,624
퍼지	0	0	10,182	8,941	1,183

산출한 것이다.

다음 설명은 검증지점 2에서의 영상소별 분류 결과가 다르게 나온 영상소를 중심으로 그 이유에 대한 설명을 정리하여 놓은 것이다.

검증지점 2에서 (1)번 영상소의 경우 항공사진에서는 논과 도로가 비슷하게 분포하고 있다. 최대우도분류법과 퍼지분류법은 농경목초지로 분류하였고 신경망분류법의 경우 인공구조물로 분류하였다. 퍼지분류법의 MGV를 살펴보면 농경목초지가 100%를 차지하는 것으로 나타나 있어 항공사진과는 대조를 이루고 있다.

(5)번은 도로변의 인도블럭과 주택지역으로 분포하고 있다. 최대우도분류법과 신경망분류법에서는 인공구조물로 분류하였으며, 퍼지분류법은 인공구조물 54%와 나대지 46%의 혼합으로 분류하였다.

(19)번 영상소의 경우 항공사진상에는 농경목초지와 도로 등의 인공구조물로 구성되어 있으나 세 가지 방법 모두 인공구조물로 분류하였다.

(20)번 영상소의 경우 항공사진 분석결과 나대지인 도로변이 대부분을 차지하고 있다. 최대우도분류법과 신경망분류법은 나대지로 분류하였으며, 퍼지분류법에서는 MGV가 나대지 0.2877, 인공구조물이 0.7177로 나타났다.

5. 결 론

이 연구에서는 춘천시 일부 지역을 연구대상으로 하여 1996년 9월 1일에 취득된 Landsat TM 영상을 이용하여 최대우도분류법, 신경망분류법, 퍼지분류법을 통하여 토지피복분류를 실시한 후, 1997년 11월 9일 촬영된 항공사진을 이용하여 분류정확도에 대한 검증을 실시함으로써 낮은 분류정확도 및 영상소내에서의 분류항목의 분포 등에 대한 일반적 토지피복분류의 연구에서

의 문제점을 해결하고자 하였으며, 다음과 같은 결과를 얻을 수 있었다.

1. 지상의 토지피복 분류항목이 비교적 균일하게 구성 되어 있는 경우 세 가지 분류기법 모두 비슷한 결과를 나타냈다. 즉 현재 많이 사용하고 있는 최대우도분류법을 이용해도 단일항목으로 분포된 면적이 많은 경우에는 같은 결과를 나타내므로 처리 시간, 처리비용 등을 감안 하였을 경우 다른 기법을 사용함으로써 생기는 컴퓨팅 비용을 줄일 수 있었다.

2. 분류항목이 혼재되어 있거나 분류항목간의 경계인 곳에서는 신경망분류법이 다소 좋은 결과를 나타냄을 알 수 있었으며, 특히 물과 인공구조물의 경계설정에 있어서 통계학적인 분류방법보다 분류정확도를 높일 수 있음을 알 수 있었다.

3. 한 영상소내에 포함되어 있을 가능성이 있는 여러 가지 분류항목에 대한 MGVI를 계산함으로써 분류 결과를 결정하는 퍼지분류법을 이용하면 여러 가지 분류항목이 복잡하게 혼재되어 있는 지역에 대한 효율적인 주제 도 제작과 GIS 모델링용 토지피복분류 영상의 제작이 가능할 것이다.

4. 신경망 이론과 퍼지 집합이론을 이용한 원격탐사 영상의 토지피복 분류결과를 수문/환경/농업/삼림/도시계획/자원관리 등 여러 GIS 응용분야에 이용할 수 있다면 보다 정확하고 개관적인 모델링 및 분석결과를 얻을 수 있을 것이다.

5. 세 가지 분류법에 의한 검증방법으로써 항공사진을 기하보정하여 이용하였으나 계절적인 차이와 시간적인 차이로 인하여 원격탐사 영상 촬영당시의 토지피복의 상

태를 정확히 파악하기 힘들었으며, 최대한 현재와 같은 상태라고 추정된 지역을 검증지역으로 선정하였다. 차후에 대한 연구과제로서 촬영시기가 같은 원격탐사 영상과 항공사진을 이용한다면 지금보다는 정확한 검증을 실시할 수 있을 것이라 생각되며, 신경망 이론과 퍼지 집합 이론을 접목한다면 두 가지의 장점을 모두 적용한 새로운 원격탐사 영상분류 기법을 개발할 수 있을 것이라 생각된다.

참고문헌

1. 김대수, "신경망의 이론과 응용 I", 하이테크정보, 1999, pp. 283-298.
2. 김대수, "신경망의 이론과 응용 II", 하이테크정보, 1999, pp. 190-201.
3. 김홍규, "퍼지집합 이론을 이용한 원격탐사 영상의 최대우도 토지피복 분류 및 변화탐지 기법의 개발", 박사학위논문, 강원대학교 일반대학원 토목공학과, 1999.
4. 오창석, 뉴로컴퓨터, 지성출판사, 1996, pp.5.
5. 양인태, 김홍규, 신계중, "다시기 원격탐사 영상의 퍼지 감독분류 결과를 이용한 토지피복 변화탐지기법의 개발", 대한토목학회논문집, 제19권, 제III-4호, 1999, pp. 687-694.
6. Cihan H. Dagli, "Artificial Neural Networks for Intelligent Manufacturing", Chapman & Hall, 1994, pp. 413-434.
7. John R. Jensen, "Introductory Digital Image Processing", 2 Edition, Prentice Hall, 1996, pp. 240-243.

(2002년 2월 5일 원고접수)