

# 소프트 컴퓨팅기술을 이용한 원격탐사 다중 분광 이미지 데이터의 분류에 관한 연구\*

—Rough 집합을 중심으로—

원 성 현\*\*

〈목 차〉

I. 서론	3. 실험 데이터들의 분포 특성
II. Rough 집합이론을 이용한 다중 분광 이미지 데이터의 밴드 분류	4. Rough 집합이론을 이용한 밴드 분류
1. Rough 집합이론	5. 특수한 분포를 갖는 데이터들의 밴드 분류
2. Rough 집합이론을 이용한 밴드 분류	6. 실험 결과의 평가
III. 실험 및 결과의 고찰	V. 결론
1. 실험대상 선정	참고문헌
2. 학습 데이터들의 분포 특성	Abstract

## I. 서 론

최근 인공위성으로부터 수집한 원격탐사 이미지 데이터(remote sensed image data)를 처리하여 실에 활용하고자 하는 노력이 사회 각 분야에서 폭 넓게 진행되고 있다. 원격탐사 데이터에 대한 구체적인 활용은 지리정보시스템(Geographic Information System, 이후 GIS로 약칭) 등으로 귀결될 수 있는데, 이러한 GIS를 통하여 효과적인 환경감시, 국토개발, 자원탐사, 기상예측, 군사동향 파악 및 농업 생산량 추정 등을 수행할 수 있다(김진일, 1994).

\* 이 논문은 1997년 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음.

\*\* 지산대학 전자계산과 조교수

원격 탐사 이미지 데이터의 종류로는 항공사진과 위성사진 등 여러 가지가 있을 수 있는데 여기서는 그 중 인공위성에서 탐사된 다중 분광 이미지 데이터(multispectral image data)를 지칭하는 것으로 한다(양영규, 1991).

다중 분광 영상 데이터에 대한 처리 과정에서 발생할 수 있는 주된 연구 분야는, 첫째 이들 이미지 데이터의 효과적인 처리라고 할 수 있다. 다중 분광 이미지 데이터 처리 시스템은 파장대에 따라 나뉘어진 여러 개의 밴드로부터 수집된 데이터를 처리하고 있다. 처리에서 가장 중요하게 인식되는 과정은 다시 두 가지 측면에서 생각할 수 있는데, 첫째는 분류(classification)라고 할 수 있다. 분류란 어떤 대상이 되는 객체(혹은 화소(pixel))가 있을 때, 그 객체가 어떤 범주(category)에 해당되는지를 결정하여 그 범주로 귀속시키는 것이다(Abe, 1997).

또 하나의 중요한 문제는 하나의 객체가 두 개 이상의 범주로 분류되었을 때, 이에 대한 정확한 성분 비율을 추정하는 문제이다. 모든 다중 분광 이미지 데이터는 하나의 화소로 표현되는 실제 지형공간의 범위가 20m×20m 혹은 30m×30m 등이다. 따라서, 이 공간내에 두 개 이상의 범주가 혼합되어 있을 때(흔히 이 경우를 혼합 화소(MIXed pixEL : MIXEL)라 한다) 이들의 성분 비율을 정확하게 추정하는 것이 매우 중요한 연구 과제가 되고 있다(Canters, 1997). 두 번째 연구 분야는 최근에 부각된 것으로 다중 밴드(multiple band)에서의 유효 밴드 선택에 관한 문제이다. 이것을 학자들에 따라서는 밴드 특징 추출 혹은 밴드 분류라고도 한다. 다중 밴드의 차원이 4차원 혹은 7차원 등이었을 때, 실제 시스템이 처리하는 밴드는 이들 전부가 아니고 그 중 의미있는 2~3개의 밴드만을 추출한다. 이때 기준에는 파장대의 물리적 특성이나 통계적 분리 가능성(separability)에 따라 밴드를 선택하지만 통상 경험에 의한 방법을 사용하고 있다. 즉, 임의의 밴드를 선택·조합하여 그 중 가장 양호한 분석 정확도를 나타내는 밴드를 조합하게 된다. 그러나, 이러한 방법은 최근 다중 밴드의 차원이 220개 등으로 확장되면서 이러한 전통적인 방법에 의한 밴드 선택에 문제가 있음이 밝혀졌다. 그것은 220여 개의 밴드에서 경험적인 방법을 사용하기에는 많은 시행착오가 발생한다는 것이다. 따라서, 수많은 분광 밴드에서 사용 가능한 유효 밴드 선택 문제가 새로운 연구 과제로 부각되고 있다(Jimenez, 1998).

본 논문에서는 이와 같이 다중 분광 이미지 데이터 처리에서 발생할 수 있는 몇 가지 연구과제 중에서 밴드 분류 문제에 대해 집중적으로 연구하기로 한다. 아직 7개의 밴드를 넘는 초다중 분광(hyperspectral) 데이터에 대한 일반적인 처리는 이루어지지 않고 있으나 향후, 이것이 보편화될 것으로 예측되고 나아가서는 220밴드 이상의 초다중

분광 밴드로 확장될 가능성이 매우 높은 현 시점에서 이에 대비한 데이터 분포 특성 분석 중심의 밴드 분류 알고리즘을 개발하는 것은 매우 중요한 일이다.

다중 분광 밴드의 분류에 대한 연구가 부각된 것은 1998년 Jimenez에 의해서였다. 그러나, 그의 방법은 지나친 연산 오버헤드를 유발하는 것이 문제점으로 지적되었고, 이 오버헤드를 Rough 집합이론의 식별 가능성을 이용하여 해결하려는 노력이 있었으나 실험적으로 그 타당성을 입증하지 못했다(원성현, 1997:1998). 따라서 본 논문에서는 충분한 실험을 통해서 그 타당성을 입증하고자 한다.

## II. Rough 집합이론을 이용한 다중 분광 이미지 데이터의 밴드 분류

### 1. Rough 집합이론

#### 1.1 Rough 집합이론의 특징

Rough 집합이론은 1982년 폴란드의 과학자 Pawlak에 의해 제안된 방법으로 객체들의 속성값들의 특징을 분석하여 객체들을 군집화하는 방법으로 많이 사용되는 이론이다(Pawlak, 1982: 1991). Rough 집합이론은 불충분하거나 일관성이 없는 정보로부터 체계적이고 자동적으로 정보를 가려냄으로써 여러 속성값들로부터 구분화 혹은 근사화할 수 있다. Rough 집합개념은 모호성과 불확실성에 대한 새로운 수학적 접근 방법이고 우리가 관심을 갖는 어떤 정보를 구성하는 전체 집합의 모든 객체에 대한 가정으로부터 출발한다. 예를 들어, 어떤 질병으로부터 고통받는 환자들의 집합을 객체들의 집합이라 하고, 질병의 징후를 환자에 대한 정보 유형이라고 할 수 있다. 동일한 정보에 의해 특징지워지는 객체는 그들에 대한 사용 가능한 정보라는 관점에서 식별 불가능하다고 본다. 이같은 방법으로 생성된 식별 불능 관계가 Rough 집합이론의 수학적 기초이다. 모든 식별 불능한 객체에 대한 어떤 집합을 기본집합(elementary set)이라 하고, 이는 전체에 대한 지식의 기초 입자 유형이라 할 수 있다. 모든 Rough 집합은 경계선을 갖지만 일반 집합(crisp set)은 그렇지 않다는 특징이 있다.

Rough 집합이론을 이용한 시스템으로부터 얻을 수 있는 장점은 다음과 같다.

첫째, 데이터 속에 숨겨진 패턴을 찾기 위한 효과적인 알고리즘을 개발하기 용이하다.

둘째, 데이터 속에 존재하는 패턴 특성의 중복을 제거하여 최소 집합화하기 용이하다. 셋째, 데이터의 의미나 중요성을 평가할 수 있다. 넷째, 데이터로부터 결정 규칙 집합을 생성하기 용이하다. 다섯째, 처리 과정을 이해하기 용이하다. 여섯째, 얻어진 결과를 해석하는 것이 간단하다. 마지막으로, 병렬처리 시스템과 같은 응용 분야에 특히 유용하다.

이들 특징의 대부분이 데이터들의 특징에 대한 분석 혹은 평가와 관련된 사항들이다. 따라서, Rough 집합이론을 이용하면 불완전한 데이터들로부터 효과적인 처리가 가능하고, 특히, 패턴 특성의 분류, 즉 군집화에 매우 유용하다는 특징으로 귀결된다.

본 논문에서도 Rough 집합이론의 이와 같은 특징 및 장점을 이용하여 다중 분광 이미징 데이터의 밴드를 분류할 것이고 이로부터 밴드 분류의 자동화를 추구할 수 있음을 보인다.

## 1.2 Rough 집합이론의 기본 개념

Rough 집합이론을 구성하는 몇 가지 정의를 먼저 소개하면 다음과 같다.

다음과 같은 속성 집합  $Q$ 가 있고 그 원소로  $n$ 개의 각 속성들이 있다고 가정하자.

$$Q = \{q_1, q_2, \dots, q_n\} \quad \dots\dots(1)$$

그리고, 분류대상이 되는  $m$ 개의 객체에 대한 집합  $X$ 와 그들의 원소  $x_1, x_2, \dots, x_m$ 이 있다고 한다.

$$X = \{x_1, x_2, \dots, x_m\} \quad \dots\dots(2)$$

또한, 속성  $q_j$ 가 보일 수 있는 속성값을 다음과 같이 표현하기로 하고 이 집합을  $V_{q_j}$ 라 하자.

$$V_{q_j} = \{\alpha, \beta, \dots, \omega\} \quad (\text{여기서, } j = 1, 2, \dots, n) \quad \dots\dots(3)$$

그러면, 이들에 대한 속성값 기술 함수  $P_X$ 는 다음과 같다.

(정의 1) 속성값 기술 함수  $P_X$

$$PX : Q \rightarrow V : P_X(q_i) = P(X, q_i) \quad \dots\dots(4)$$

(여기서,  $Q \rightarrow V$ 는 속성 집합  $Q$ 로부터 속성값 집합  $V$ 로의 사상)

그리고, 두 개의 객체 집합이 임의의 속성에 대해 식별할 수 없음을 나타내는 식별 불능 관계  $\text{ind}(Q)$ 는 다음과 같이 정의한다.

(정의 2) 식별 불능 관계(indiscernibility relation)  $\text{ind}(Q)$

객체  $x_i$ 와  $x_j$ 가 어떤 속성  $q_i$ 에 관하여 식별 불능 관계라면 다음과 같이 나타낸다.

$$\begin{aligned} P(x_i, q_i) &= P(x_j, q_i) \\ (x_i, x_j) &\in \text{ind}(q_i) \\ R &= \text{ind}(Q) \end{aligned} \quad \dots\dots(5)$$

(여기서,  $R$ 은 두 개의 객체집합  $x_i$ 와  $x_j$ 가 상호 식별 불능 관계인 동치 관계임)

그러므로,  $x_i, x_j$ 가  $P \subset Q$ 에 의하여 식별될 수 없을 경우에 다음과 같이 나타낸다.

$$\begin{aligned} (x_i, x_j) &\in \text{ind}(P) \\ \text{ind}(P) &= \bigcap_{q_i \in P} \text{ind}(q_i) \end{aligned} \quad \dots\dots(6)$$

여기서,  $P=Q$ 일 때,  $(x_i, x_j) \in \text{ind}(q_i)$ 라면  $x_i$ 와  $x_j$ 는 상호 식별될 수 없는 관계가 되고,  $\text{ind}(q_i)$ 는 동치 관계이므로 이것에 의한 분할은 상집합(quotient set)이 된다.

$$X/\text{ind}(q_i) = \{[x_i] \mid x_i \in X\} \quad \dots\dots(7)$$

## 2. Rough 집합이론을 이용한 밴드 분류

본 절에서는 앞에서 소개한 이론을 바탕으로 Rough 집합을 이용한 동치 클래스 구성 방법에 대해 설명한다. 먼저, 일반적인 동치 클래스 구성을 위한 가정과 표기법을 소개하고, 이를 통해 그 수행 절차를 설명한 후, 예를 통해 확인한다.

## 2.1 Rough 집합이론을 이용한 동치 클래스 구성

Rough 집합이론의 식별 불능 관계를 이용하여 객체들을 분류하기 위해서는 먼저 속성 집합과 객체 집합이 정의되어 있어야 한다. 속성집합  $Q$ 와 객체집합  $X$ 가 정의되어 있다고 하자. 그러면 객체  $x_i$ 가 취하는 속성값  $AV(x_i)$ 는 다음과 같이 속성  $q_j$ 가 가질 수 있는 값 중의 한 원소로 주어지게 된다.

$$q_j\{AV(x_i)\} = \text{one of } \{\alpha, \beta, \dots, \omega\} \quad \dots\dots(8)$$

따라서, 속성  $q_k$ 에 대한 식별 불능 관계에 해당되는 객체는 다음 식에 의해 구할 수 있다.

$$X/\text{ind}(q_k) = \{x_i, x_j \mid q_k\{AV(x_i)\} = q_k\{AV(x_j)\}\} \quad \dots\dots(9)$$

(여기서,  $i, j = 1, 2, \dots, m$ )

만일, 분류기준이 되는 속성이  $q_j, q_k$ 로 두 개라면 다음과 같이 식별 불능 관계에 해당되는 객체를 구할 수 있다.

$$X/\text{ind}(q_j, q_k) = \{x_i, x_j \mid (q_j, q_k)\{AV(x_i)\} = (q_j, q_k)\{AV(x_j)\}\} \quad \dots\dots(10)$$

(여기서,  $i, j = 1, 2, \dots, m$ )

모든 속성  $Q$ 에 대해 식별 불능 관계에 해당되는 객체는 다음 식에 의해 구할 수 있다.

$$X/\text{ind}(Q) = \{x_i, x_j \mid Q\{AV(x_i)\} = Q\{AV(x_j)\}\} \quad \dots\dots(11)$$

(여기서,  $i, j = 1, 2, \dots, m$ )

## 2.2 동치 클래스 구성 예

다음과 같은 정보시스템 테이블이 구성되어 있다고 가정하자.  $x_i$ 는 환자이고,  $q_i$ 는 검사 항목,  $V$ 는 검사 결과의 집합이라고 한다면 다음과 같이 표현할 수 있다.

$$X = \{x_1, x_2, x_3, x_4, x_5\} \quad \dots\dots(12)$$

$$Q = \{q_1, q_2, q_3, q_4\} \quad \dots\dots(13)$$

$$V = \{0, 1, 2\} \quad \dots\dots(14)$$

〈표 1〉은 환자들의 검사항목에 대한 검사 결과를 나타내는 표이다. 이로부터 검사의 결과는 0, 1, 2의 세 가지 종류로 얻어졌음을 알 수 있다.

〈표 1〉 환자들의 임상 검사 항목에 대한 정보 시스템 테이블

검사항목(Q) 환자(X)	q <sub>1</sub>	q <sub>2</sub>	q <sub>3</sub>	q <sub>4</sub>
x <sub>1</sub>	1	1	0	1
x <sub>2</sub>	1	1	0	1
x <sub>3</sub>	1	1	1	0
x <sub>4</sub>	0	1	1	2
x <sub>5</sub>	0	1	1	1

이 정보시스템 테이블을 Rough 집합의 동치 관계를 이용하여 동치 클래스화하면 다음과 같은 분류가 가능하다.

$$X/\text{ind}(q_1) = \{[x_1, x_2, x_3], [x_4, x_5]\} \quad \dots\dots(15)$$

$$X/\text{ind}(q_2) = \{[x_1, x_2, x_3, x_4, x_5]\} \quad \dots\dots(16)$$

$$X/\text{ind}(q_3) = \{[x_1, x_2], [x_3, x_4, x_5]\} \quad \dots\dots(17)$$

$$X/\text{ind}(q_4) = \{[x_1, x_2, x_5], [x_3], [x_4]\} \quad \dots\dots(18)$$

$$X/\text{ind}(q_1, q_2) = \{[x_1, x_2, x_3], [x_4, x_5]\} \quad \dots\dots(19)$$

$$X/\text{ind}(Q) = \{[x_1, x_2], [x_3], [x_4], [x_5]\} \quad \dots\dots(20)$$

즉, 식 (15)에 의하면 q<sub>1</sub>이라는 속성값에 의해 x<sub>1</sub>, x<sub>2</sub>, x<sub>3</sub>은 식별 불능 관계, 다시 말하면 조합 가능 관계이지만 이들과 x<sub>4</sub>, x<sub>5</sub>는 식별 가능 관계, 즉 조합 불가능 관계라고 말할 수 있다. 또한, 식 (19)에 의하면 q<sub>1</sub>과 q<sub>2</sub>라는 두 개의 속성값에 의해 x<sub>1</sub>, x<sub>2</sub>, x<sub>3</sub>은 식별 불능 관계이고 x<sub>4</sub>, x<sub>5</sub> 역시 상호 식별 불능 관계이다. 그러나, 이들 두 군집 간에는 식별 가능 관계이다. 식 (20)에 의하면 속성 집합 Q, 즉 모든 속성들에 대해 x<sub>1</sub>과 x<sub>2</sub>는

상호 식별 불능 관계, 즉  $P(x_1, Q) = P(x_2, Q)$ 이다. 그러나, 이들과 다른 검사 항목  $x_3, x_4, x_5$ 들 사이에는 식별 가능 관계가 성립된다.

## 2.3 Rough 집합이론을 이용한 다중 분광 이미지 데이터의 밴드 분류

### 2.3.1 다중 분광 영상 데이터의 참조표

통상적으로 임의의 대역에서 특정 범주에 속하는 화소들의 분광 강도는 거의 정해져 있다. 물론, 데이터 수집과정에서 왜곡 현상이 발생할 수 있고, 의미없는 값(garbage)들이 수집되는 경우도 있지만 이들은 이미 개발되어 있는 여러 보정 방법들에 의해서 보정되므로 정해져 있는 분광 강도에서의 표준편차는 그리 크지 않다. 그러나, 어느 정도의 편차는 발생할 수 있기 때문에 다음과 같은 일반화된 다중 분광 영상 데이터의 참조표(look up table)를 만들어 분류시 활용하기로 한다. 이 참조표는 많은 데이터들을 수차례의 학습을 시킨 결과로부터 만들어진 것이다.

<그림 1>은  $n$ 개의 면(frame)에 대한 범주와 밴드간의 관계를 나타내는 분광 강도의 참조표이다. 각 면에서  $C$ 는 범주들의 집합으로  $C_1$ 부터  $C_m$ 까지의  $m$ 개로 구성되어 있다. 즉, 분류하고자 하는 대상들이 귀속될 수 있는 범주를 의미하는데 예를 들면, 숲, 수역 등이다. 또한,  $B$ 는 밴드들의 집합으로  $B_1$ 부터  $B_7$ 까지의 7개의 밴드로 설정했다. 본 논문의 목표는 초다중 분광 밴드 환경에서의 특징 추출이지만 그 근거는 7개의 밴드 환경인 다중 분광 밴드에서 만들하고자 한다.  $v_{111}, v_{27m}$  등은 특정 면( $v_{111}$ 은 1면,  $v_{27m}$ 은 2면)의 특정 밴드( $v_{111}$ 은 밴드 1,  $v_{27m}$ 은 밴드 7)에서 특정 범주( $v_{111}$ 은 범주 1,  $v_{27m}$ 은 범주  $m$ )로 귀속되는 화소가 갖는 분광 강도이다.

<그림 2>는 이들  $n$ 개의 면에 대해 각 면의 동일한 범주와 동일한 밴드에서 발생하는 다양한 분광 강도에 대한 평균( $\mu$ )을 구하고 이에 대한 표준편차( $\sigma$ )를 구하여 오차의 한계로 인정한 것이다.  $\mu_{11}$ 은  $n$ 개의 면에 대하여 밴드 1( $B_1$ )에서 범주 1( $C_1$ )을 나타내는 분광 강도들의 평균을 의미하고,  $\sigma_{11}$ 은 밴드 1에서 범주  $C_1$ 로 분류되는  $n$ 개의 면의 분광 강도에 대한 표준편차를 의미한다.



<p>frame 1</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="border: none;"></td> <td style="border: none;"><b>C</b></td> <td style="border: none;"></td> <td style="border: none;"></td> <td style="border: none;"></td> <td style="border: none;"></td> </tr> <tr> <td style="border: none;"><b>B</b></td> <td style="border: none;"></td> <td style="border: none;"><math>C_1</math></td> <td style="border: none;"><math>C_2</math></td> <td style="border: none;"><math>\dots</math></td> <td style="border: none;"><math>C_m</math></td> </tr> <tr> <td style="border: none;"><math>B_1</math></td> <td style="border: none;"></td> <td><math>v_{111}</math></td> <td><math>v_{112}</math></td> <td><math>\dots</math></td> <td><math>v_{11m}</math></td> </tr> <tr> <td style="border: none;"><math>B_2</math></td> <td style="border: none;"></td> <td><math>v_{121}</math></td> <td><math>v_{122}</math></td> <td><math>\dots</math></td> <td><math>v_{12m}</math></td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;"><math>B_7</math></td> <td style="border: none;"></td> <td><math>v_{171}</math></td> <td><math>v_{172}</math></td> <td><math>\dots</math></td> <td><math>v_{17m}</math></td> </tr> </table>		<b>C</b>					<b>B</b>		$C_1$	$C_2$	$\dots$	$C_m$	$B_1$		$v_{111}$	$v_{112}$	$\dots$	$v_{11m}$	$B_2$		$v_{121}$	$v_{122}$	$\dots$	$v_{12m}$	.		.	.	$\dots$	.	.		.	.	$\dots$	.	.		.	.	$\dots$	.	$B_7$		$v_{171}$	$v_{172}$	$\dots$	$v_{17m}$	<p>frame 2</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="border: none;"></td> <td style="border: none;"><b>C</b></td> <td style="border: none;"></td> <td style="border: none;"></td> <td style="border: none;"></td> <td style="border: none;"></td> </tr> <tr> <td style="border: none;"><b>B</b></td> <td style="border: none;"></td> <td style="border: none;"><math>C_1</math></td> <td style="border: none;"><math>C_2</math></td> <td style="border: none;"><math>\dots</math></td> <td style="border: none;"><math>C_m</math></td> </tr> <tr> <td style="border: none;"><math>B_1</math></td> <td style="border: none;"></td> <td><math>v_{211}</math></td> <td><math>v_{212}</math></td> <td><math>\dots</math></td> <td><math>v_{21m}</math></td> </tr> <tr> <td style="border: none;"><math>B_2</math></td> <td style="border: none;"></td> <td><math>v_{221}</math></td> <td><math>v_{222}</math></td> <td><math>\dots</math></td> <td><math>v_{22m}</math></td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;"><math>B_7</math></td> <td style="border: none;"></td> <td><math>v_{271}</math></td> <td><math>v_{272}</math></td> <td><math>\dots</math></td> <td><math>v_{27m}</math></td> </tr> </table> <p style="text-align: right;">...</p>		<b>C</b>					<b>B</b>		$C_1$	$C_2$	$\dots$	$C_m$	$B_1$		$v_{211}$	$v_{212}$	$\dots$	$v_{21m}$	$B_2$		$v_{221}$	$v_{222}$	$\dots$	$v_{22m}$	.		.	.	$\dots$	.	.		.	.	$\dots$	.	.		.	.	$\dots$	.	$B_7$		$v_{271}$	$v_{272}$	$\dots$	$v_{27m}$
	<b>C</b>																																																																																																
<b>B</b>		$C_1$	$C_2$	$\dots$	$C_m$																																																																																												
$B_1$		$v_{111}$	$v_{112}$	$\dots$	$v_{11m}$																																																																																												
$B_2$		$v_{121}$	$v_{122}$	$\dots$	$v_{12m}$																																																																																												
.		.	.	$\dots$	.																																																																																												
.		.	.	$\dots$	.																																																																																												
.		.	.	$\dots$	.																																																																																												
$B_7$		$v_{171}$	$v_{172}$	$\dots$	$v_{17m}$																																																																																												
	<b>C</b>																																																																																																
<b>B</b>		$C_1$	$C_2$	$\dots$	$C_m$																																																																																												
$B_1$		$v_{211}$	$v_{212}$	$\dots$	$v_{21m}$																																																																																												
$B_2$		$v_{221}$	$v_{222}$	$\dots$	$v_{22m}$																																																																																												
.		.	.	$\dots$	.																																																																																												
.		.	.	$\dots$	.																																																																																												
.		.	.	$\dots$	.																																																																																												
$B_7$		$v_{271}$	$v_{272}$	$\dots$	$v_{27m}$																																																																																												
<p>frame n-1</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="border: none;"></td> <td style="border: none;"><b>C</b></td> <td style="border: none;"></td> <td style="border: none;"></td> <td style="border: none;"></td> <td style="border: none;"></td> </tr> <tr> <td style="border: none;"><b>B</b></td> <td style="border: none;"></td> <td style="border: none;"><math>C_1</math></td> <td style="border: none;"><math>C_2</math></td> <td style="border: none;"><math>\dots</math></td> <td style="border: none;"><math>C_m</math></td> </tr> <tr> <td style="border: none;"><math>B_1</math></td> <td style="border: none;"></td> <td><math>v_{n-111}</math></td> <td><math>v_{n-112}</math></td> <td><math>\dots</math></td> <td><math>v_{n-11m}</math></td> </tr> <tr> <td style="border: none;"><math>B_2</math></td> <td style="border: none;"></td> <td><math>v_{n-121}</math></td> <td><math>v_{n-122}</math></td> <td><math>\dots</math></td> <td><math>v_{n-12m}</math></td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;"><math>B_7</math></td> <td style="border: none;"></td> <td><math>v_{n-171}</math></td> <td><math>v_{n-172}</math></td> <td><math>\dots</math></td> <td><math>v_{n-17m}</math></td> </tr> </table>		<b>C</b>					<b>B</b>		$C_1$	$C_2$	$\dots$	$C_m$	$B_1$		$v_{n-111}$	$v_{n-112}$	$\dots$	$v_{n-11m}$	$B_2$		$v_{n-121}$	$v_{n-122}$	$\dots$	$v_{n-12m}$	.		.	.	$\dots$	.	.		.	.	$\dots$	.	.		.	.	$\dots$	.	$B_7$		$v_{n-171}$	$v_{n-172}$	$\dots$	$v_{n-17m}$	<p>frame n</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="border: none;"></td> <td style="border: none;"><b>C</b></td> <td style="border: none;"></td> <td style="border: none;"></td> <td style="border: none;"></td> <td style="border: none;"></td> </tr> <tr> <td style="border: none;"><b>B</b></td> <td style="border: none;"></td> <td style="border: none;"><math>C_1</math></td> <td style="border: none;"><math>C_2</math></td> <td style="border: none;"><math>\dots</math></td> <td style="border: none;"><math>C_m</math></td> </tr> <tr> <td style="border: none;"><math>B_1</math></td> <td style="border: none;"></td> <td><math>v_{n11}</math></td> <td><math>v_{n12}</math></td> <td><math>\dots</math></td> <td><math>v_{n1m}</math></td> </tr> <tr> <td style="border: none;"><math>B_2</math></td> <td style="border: none;"></td> <td><math>v_{n21}</math></td> <td><math>v_{n22}</math></td> <td><math>\dots</math></td> <td><math>v_{n2m}</math></td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;">.</td> <td style="border: none;"></td> <td>.</td> <td>.</td> <td><math>\dots</math></td> <td>.</td> </tr> <tr> <td style="border: none;"><math>B_7</math></td> <td style="border: none;"></td> <td><math>v_{n71}</math></td> <td><math>v_{n72}</math></td> <td><math>\dots</math></td> <td><math>v_{n7m}</math></td> </tr> </table>		<b>C</b>					<b>B</b>		$C_1$	$C_2$	$\dots$	$C_m$	$B_1$		$v_{n11}$	$v_{n12}$	$\dots$	$v_{n1m}$	$B_2$		$v_{n21}$	$v_{n22}$	$\dots$	$v_{n2m}$	.		.	.	$\dots$	.	.		.	.	$\dots$	.	.		.	.	$\dots$	.	$B_7$		$v_{n71}$	$v_{n72}$	$\dots$	$v_{n7m}$
	<b>C</b>																																																																																																
<b>B</b>		$C_1$	$C_2$	$\dots$	$C_m$																																																																																												
$B_1$		$v_{n-111}$	$v_{n-112}$	$\dots$	$v_{n-11m}$																																																																																												
$B_2$		$v_{n-121}$	$v_{n-122}$	$\dots$	$v_{n-12m}$																																																																																												
.		.	.	$\dots$	.																																																																																												
.		.	.	$\dots$	.																																																																																												
.		.	.	$\dots$	.																																																																																												
$B_7$		$v_{n-171}$	$v_{n-172}$	$\dots$	$v_{n-17m}$																																																																																												
	<b>C</b>																																																																																																
<b>B</b>		$C_1$	$C_2$	$\dots$	$C_m$																																																																																												
$B_1$		$v_{n11}$	$v_{n12}$	$\dots$	$v_{n1m}$																																																																																												
$B_2$		$v_{n21}$	$v_{n22}$	$\dots$	$v_{n2m}$																																																																																												
.		.	.	$\dots$	.																																																																																												
.		.	.	$\dots$	.																																																																																												
.		.	.	$\dots$	.																																																																																												
$B_7$		$v_{n71}$	$v_{n72}$	$\dots$	$v_{n7m}$																																																																																												

단, 범주들의 집합  $C = \{C_1, C_2, \dots, C_m\}$   
 대역들의 집합  $B = \{B_1, B_2, \dots, B_7\}$   
 $v_{ijk} = i$ 번째 frame의  $j$ 번째 대역에서 범주  $k$ 로 귀속될 학습된 분광 강도

<그림 1> 다중 분광 이미지 데이터의 참조표

	<b>C</b>				
<b>B</b>		$C_1$	$C_2$	$\dots$	$C_m$
$B_1$		$[\mu_{11}^- - \sigma_{11}, \mu_{11}^+ + \sigma_{11}]$	$[\mu_{12}^- - \sigma_{12}, \mu_{12}^+ + \sigma_{12}]$	$\dots$	$[\mu_{1m}^- - \sigma_{1m}, \mu_{1m}^+ + \sigma_{1m}]$
$B_2$		$[\mu_{21}^- - \sigma_{21}, \mu_{21}^+ + \sigma_{21}]$	$[\mu_{22}^- - \sigma_{22}, \mu_{22}^+ + \sigma_{22}]$	$\dots$	$[\mu_{2m}^- - \sigma_{2m}, \mu_{2m}^+ + \sigma_{2m}]$
.		.	.	$\dots$	.
.		.	.	$\dots$	.
.		.	.	$\dots$	.
$B_7$		$[\mu_{71}^- - \sigma_{71}, \mu_{71}^+ + \sigma_{71}]$	$[\mu_{72}^- - \sigma_{72}, \mu_{72}^+ + \sigma_{72}]$	$\dots$	$[\mu_{7m}^- - \sigma_{7m}, \mu_{7m}^+ + \sigma_{7m}]$

단,  $\mu_{ij} = n$ 개의 frame에 대하여 대역  $i$ 에서 범주  $j$ 로 귀속될 학습된 분광 강도의 평균  
 $\sigma_{ij} = n$ 개의 frame에 대하여 대역  $i$ 에서 범주  $j$ 로 귀속될 학습된 분광 강도의 표준편차

<그림 2> 다중 분광 이미지 데이터에 허용된 오차의 한계

### 2.3.2 탐사된 화소들의 범주 귀속

〈그림 3〉은 탐사된 객체에 대한 화소들의 분광 강도표이다. 여기서, P는 k개의 탐사 화소  $P_1, P_2, \dots, P_k$ 의 집합을 의미한다. 또한,  $p_{v11}$ 은 화소  $P_1$ 이 밴드 1에서 나타낸 분광 강도를 말한다.

B \ P	P			
	$P_1$	$P_2$	...	$P_k$
$B_1$	$p_{v11}$	$p_{v12}$	...	$p_{v1k}$
$B_2$	$p_{v21}$	$p_{v22}$	...	$p_{v2k}$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
$B_7$	$p_{v71}$	$p_{v72}$	...	$p_{v7k}$

단, 탐사된 화소들의 집합  $P = \{P_1, P_2, \dots, P_k\}$

$p_{vij}$  = 대역 i에서 화소 j가 나타낸 분광 강도

〈그림 3〉 화소들의 밴드별 분광 강도

```

for i = 1 to 7 by 1
/* 침자 i는 대역 1부터 대역 7까지를 의미 */

    for j = 1 to k by 1
/* 침자 j는 범주 귀속시키하고자 하는 화소들의 순차 번호를 의미하고 여기서는 k개로 가정 */

        for n = 1 to m by 1
/* 침자 n은 귀속될 범주의 번호로 여기서는 m개의 범주로 가정 */

            if ( $\mu_{in} - \sigma_{in} \leq p_{vij} \leq \mu_{in} + \sigma_{in}$ ) then Category =  $C_n$ 
                exit
            end
            if ( $\mu_{in} - \sigma_{in} > p_{vij}$  or  $p_{vij} > \mu_{in} + \sigma_{in}$ ) then
                for n = 1 to m by 1
                    Category =  $\min(\mu_{in}, p_{vij})$ 인  $C_n$ 
                end
            end
        end
    end
end
    
```

〈그림 4〉 범주 귀속 알고리즘

〈그림 2〉와 〈그림 3〉을 이용하여 다음 〈그림 5〉와 같이 각 화소들을 특정 범주로 귀속시킨다. 이때, 귀속되는 원칙은 화소의 분광 강도가 〈그림 2〉의 범위내에 포함되는 경우에는 해당 범주로 귀속시키고, 그렇지 않은 경우는 가장 인접한 범주로 귀속시킨다. 이에 대한 알고리즘은 〈그림 4〉와 같다.

〈그림 5〉는 〈그림 4〉의 알고리즘에 의해 탐사된 각 화소에 대해 범주 귀속시킨 결과이다.

B \ P	P			
	P <sub>1</sub>	P <sub>2</sub>	...	P <sub>k</sub>
B <sub>1</sub>	C <sub>1</sub>	C <sub>n</sub>	...	C <sub>2</sub>
B <sub>2</sub>	C <sub>2</sub>	C <sub>1</sub>	...	C <sub>1</sub>
B <sub>3</sub>	C <sub>n-1</sub>	C <sub>n</sub>	...	C <sub>1</sub>
B <sub>4</sub>	C <sub>2</sub>	C <sub>n</sub>	...	C <sub>3</sub>
B <sub>5</sub>	C <sub>2</sub>	C <sub>1</sub>	...	C <sub>2</sub>
B <sub>6</sub>	C <sub>n-1</sub>	C <sub>3</sub>	...	C <sub>n</sub>
B <sub>7</sub>	C <sub>1</sub>	C <sub>2</sub>	...	C <sub>n</sub>

〈그림 5〉 탐사된 화소들의 범주 귀속 결과

### 2.3.3 Rough 집합이론을 이용한 동치 클래스 생성 및 해석

〈그림 2〉와 〈그림 3〉을 통해 구성한 〈그림 5〉로부터 〈그림 6〉과 같이 Rough 집합이론을 이용하여 동치 클래스를 생성시킨다.

B \ P	P			
	P <sub>1</sub>	P <sub>2</sub>	...	P <sub>k</sub>
B <sub>1</sub>	C <sub>1</sub>	C <sub>n</sub>	...	C <sub>2</sub>
B <sub>2</sub>	C <sub>2</sub>	C <sub>1</sub>	...	C <sub>1</sub>
B <sub>3</sub>	C <sub>n-1</sub>	C <sub>n</sub>	...	C <sub>1</sub>
B <sub>4</sub>	C <sub>2</sub>	C <sub>n</sub>	...	C <sub>3</sub>
B <sub>5</sub>	C <sub>2</sub>	C <sub>1</sub>	...	C <sub>2</sub>
B <sub>6</sub>	C <sub>n-1</sub>	C <sub>3</sub>	...	C <sub>n</sub>
B <sub>7</sub>	C <sub>1</sub>	C <sub>2</sub>	...	C <sub>n</sub>

〈그림 6〉 탐사된 화소들의 동치 클래스(P<sub>1</sub>에 대해서만 표시)

이로부터 동치 클래스를 추출하면 다음과 같다.

$$\mathbb{B}/\text{ind}(P_1) = \{[B_1, B_7], [B_2, B_4, B_5], [B_3, B_6]\} \quad \dots\dots(21)$$

$$\mathbb{B}/\text{ind}(P_2) = \{[B_1, B_3, B_4], [B_2, B_5], [B_6, B_7]\} \quad \dots\dots(22)$$

$$\mathbb{B}/\text{ind}(P_k) = \{[B_1, B_5], [B_2, B_3], [B_4], [B_6, B_7]\} \quad \dots\dots(23)$$

$$\mathbb{B}/\text{ind}(P) = \{[B_1], [B_2], [B_3], [B_4], [B_5], [B_6], [B_7]\} \quad \dots\dots(24)$$

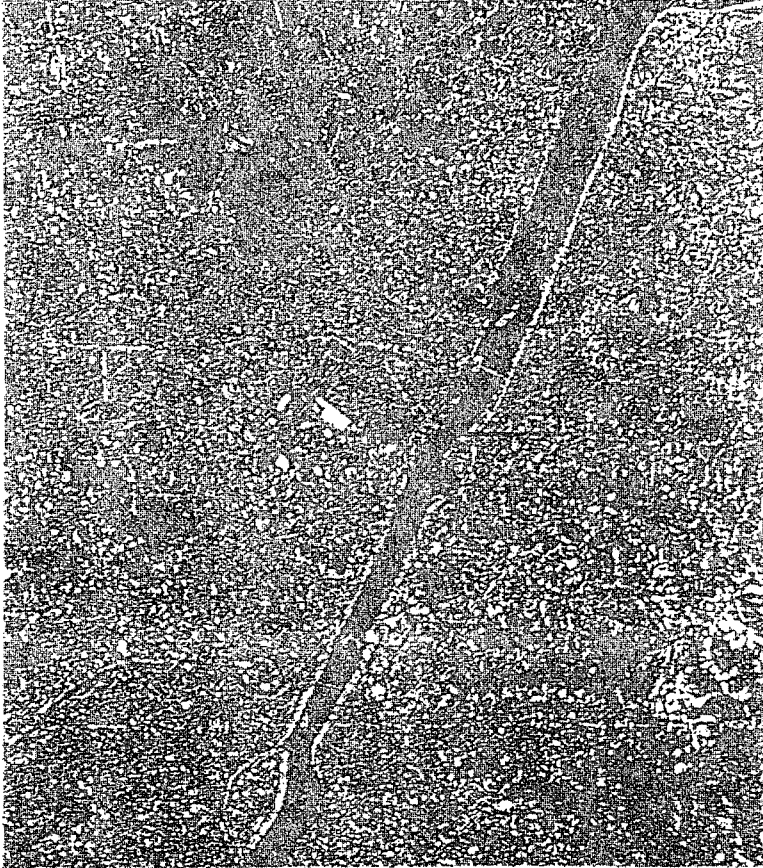
그러므로, 화소  $P_1$ 에 대해 밴드 1과 7은 식별 불능 관계이고, 밴드 2, 4, 5 역시 상호 식별 불능관계임을 나타낸다.

### Ⅲ. 실험 및 결과의 고찰

#### 1. 실험대상 선정

본 논문에서 실험대상으로 선정한 것은 1992년 6월 2일의 LANDSAT TM 데이터이고 지역은 서울의 한강과 행주대교 및 그 인근 지역이며 보조도로로 동일 지역의 국립지리원의 1 : 10000 지형도를 사용하였다. LANDSAT TM 데이터를 실험 데이터로 선정한 이유는 SPOT HRV 데이터보다 분광 해상도가 좋고 우리 나라에서는 TM 데이터가 더 구하기 쉽다는 장점이 있기 때문이다. 본래 LANDSAT TM의 장면 당 지상면적은 170km(세로)×185km(가로)이고 5,965×6,920개의 화소로 구성되어 있는데 앞서 소개한 지역에 해당되는 부분만 선택하여 136×136=18,496개의 화소로 실험을 수행했다. 여기서, LANDSAT TM 데이터의 공간 해상도가 30m×30m이기 때문에 세로로 5,965 화소×30m=178.95km의 실제공간에 해당되나 8.95km의 차이는 센서가 데이터를 수집하는 과정에서 일정 부분을 중첩하여 촬영하기 때문에 발생하는 편차이다. 이는 가로도 마찬가지이다. 또, 대상이 되는 지역에는 많은 범주가 있을 수 있지만 수역(water), 곡물(crop), 도시지역(urban), 숲(forest)의 4가지로만 제한하여 실험을 수행했다.

다음 <그림 7>은 실험대상으로 사용한 지역을 그래픽 툴인 Photoshop으로 처리하여 출력한 것이다. <그림 7>에서 사각형에 해당되는 부분이 본 논문에서 실험대상으로 사용한 지역이다.



〈그림 7〉 실험지역의 컴퓨터 출력 결과

## 2. 학습 데이터들의 분포 특성

실험에 사용될 데이터는 크게 학습 데이터와 실험 데이터로 나누어진다. 학습 데이터는 기존의 다중 분광 영상 데이터 처리 시스템과 분류 방법을 이용하여 실험 지역내의 영상 데이터를 먼저 학습하는 데 필요로 하는 데이터를 말한다. 실험 데이터는 여기서 제안하는 방법을 구현하기 위해 특정 지역을 선정하여 이들을 탐사한 데이터를 말한다.

학습 데이터로 사용하고자 하는 다중 분광 이미지 데이터의 분포 특성은 다음과 같다. 실험에 사용된  $136 \times 136$ 개의 화소 중 명확히 4개의 범주에 해당되지 않는 화소와 2개 이상의 화소에 포함될 가능성이 있는 화소들을 제외하고, 명확히 4개의 범주에 해당되는 화소로만 제한했기 때문에 전체 화소의 수는 불규칙할 수 있다.

## 2.1 학습 데이터들에 대한 분포 특성

다음은 학습 데이터들에 대한 분포 특성이다. 이들은 ERDAS라는 GIS 툴과 MLC라는 분류 알고리즘에 의해 분류된 결과를 분석한 것이다. 학습 데이터들에 대한 분포 특성을 정리하면 그 양이 너무 방대하여 여기서는 대표적인 밴드와 범주를 선택하여 정리하였다.

〈표 2〉 밴드 2에서 숲의 분포 특성

분광강도	41	42	43	44	45	46	47	48
화소 수	4	10	16	29	10	3	1	2
비중(%)	2.00	5.00	8.00	14.50	5.00	1.50	0.50	1.00
분광강도	49	50	51	52	53	74	75	76
화소 수	12	7	2	0	4	1	1	1
비중(%)	6.00	3.50	1.00	0	2.00	0.50	0.50	0.50
분광강도	77	78	79	80	81	82	83	84
화소 수	2	2	3	2	3	2	5	5
비중(%)	1.00	1.00	1.50	1.00	1.50	1.00	2.50	2.50
분광강도	85	86	87	88	89	90	91	92
화소 수	1	3	5	5	5	5	8	2
비중(%)	0.50	1.50	2.50	2.50	2.50	2.50	4.00	1.00
분광강도	93	94	95	96	97	98	99	100
화소 수	3	4	6	5	3	3	1	1
비중(%)	1.50	2.00	3.00	2.50	1.50	1.50	0.50	0.50
분광강도	101	102	103	104	105	106	107	108
화소 수	1	0	1	1	1	1	2	1
비중(%)	0.50	0	0.50	0.50	0.50	0.50	1.00	0.50
분광강도	109	110	111	계				
화소 수	3	1	1	200				
비중(%)	1.50	0.50	0.50	100				

〈표 2〉는 밴드 2에서의 숲에 대한 데이터 분포 특성이다. 분광 강도 42부터 45까지에 비교적 많은 화소가 군집되어 있는데, 그뿐 아니라 분광 강도 41부터 111까지의 거의 모든 화소들에 널리 분포되어 있다. 이와 같은 경우 표준편차가 매우 크게 발생하여 범주 귀속시키는데 오차가 발생할 가능성이 다른 경우보다 상대적으로 크다고 볼 수 있

다. 왜냐하면 본 논문에서는 데이터 분포 특성을 분석하면서 그들 데이터의 평균과 표준편차를 이용하여 일정한 범위내의 값들을 대표값으로 사용하고 있다. 그러나, 이들이 정규 분포를 이루고 있다면 큰 문제가 없으나 숲과 같이 정규분포를 나타내지 못하는 경우에는 평균을 중심으로 한 값들을 과연 대표값으로 사용할 수 있는가하는 문제가 제기된다. 정규분포인 경우에 전체의 약 68.3%에 해당되는 부분이 허용 오차내로 귀속되지만 지금과 같은 경우는 이보다 훨씬 못미치는 부분만이 허용 오차내로 귀속될 뿐이기 때문이다. 그러나, 여기서는 그를 무시하고 본 논문에서 제안하는 방법을 그대로 적용시킨다. 그리고 그 문제는 다음절에서 논의하기로 한다.

〈표 3〉은 밴드 4에서 곡물의 분포 특성을 나타낸다. 이러한 경우를 특별히 쌍봉형이라고 하는데 분광 강도 14와 52를 중심으로 두 개의 정점을 형성하며 분포되어 있다. 이와 같은 경우도 앞서 언급했듯이 범주 귀속시키기가 매우 까다로운 경우에 해당된다.

〈표 3〉 밴드 4에서 곡물의 분포 특성

분광강도	8	9	10	11	12	13	14	15	16
화소 수	1	9	13	16	6	9	17	13	7
비중(%)	0.42	3.75	5.42	6.67	2.50	3.75	7.08	5.42	2.92
분광강도	17	18	19	20	21	22	23	24	25
화소 수	7	4	4	3	1	1	1	3	3
비중(%)	2.92	1.67	1.67	1.25	0.42	0.42	0.42	1.25	1.25
분광강도	26	27	46	47	48	49	50	51	52
화소 수	0	2	2	3	7	8	11	15	19
비중(%)	0	0.83	0.83	1.25	2.92	3.33	4.58	6.25	7.92
분광강도	53	54	55	56	57	58	59	60	계
화소 수	15	13	8	11	6	1	0	1	232
비중(%)	6.25	5.42	3.33	4.58	2.50	0.42	0	0.42	100

〈표 4〉는 밴드 4에서 숲의 분포 특성을 나타내고 있는데 밴드 2에서 나타났던 특성 분포와 비슷하게 밴드 4에서도 비교적 광범위하게 분포되어 있는 것을 알 수 있다.

〈표 4〉 밴드 4에서 숲의 분포 특성

분광강도	74	75	76	77	78	79	80	81
화소 수	1	1	1	2	2	3	2	3
비중(%)	1.00	1.00	1.00	2.00	2.00	3.00	2.00	3.00
분광강도	82	83	84	85	86	87	88	89
화소 수	2	5	5	1	3	5	5	5
비중(%)	2.00	5.00	5.00	1.00	3.00	5.00	5.00	5.00
분광강도	90	91	92	93	94	95	96	97
화소 수	5	8	2	3	4	6	5	3
비중(%)	5.00	8.00	2.00	3.00	4.00	6.00	5.00	3.00
분광강도	98	99	100	101	102	103	104	105
화소 수	3	1	1	1	0	1	1	1
비중(%)	3.00	1.00	1.00	1.00	0	1.00	1.00	1.00
분광강도	106	107	108	109	110	111	계	
화소 수	1	2	1	3	1	1	100	
비중(%)	1.00	2.00	1.00	3.00	1.00	1.00	100	

〈표 5〉는 밴드 7에서의 숲을 나타내는 분포 특성표인데 역시 다양한 분광강도를 가지며 널리 분포되어 있다.

〈표 5〉 밴드 7에서 숲의 분포 특성

분광강도	21	22	23	24	25	26	27	28	29
화소 수	9	8	9	9	5	7	5	4	6
비중(%)	9.00	8.00	9.00	9.00	5.00	7.00	5.00	4.00	6.00
분광강도	30	31	32	33	34	35	36	37	38
화소 수	3	2	1	3	0	0	0	1	0
비중(%)	3.00	2.00	1.00	3.00	0	0	0	1.00	0
분광강도	39	40	41	42	43	44	45	46	47
화소 수	3	2	3	1	1	1	4	4	1
비중(%)	3.00	2.00	3.00	1.00	1.00	1.00	4.00	4.00	1.00
분광강도	48	49	50	51	52	53	계		
화소 수	1	3	1	2	0	1	100		
비중(%)	1.00	3.00	1.00	2.00	0	1.00	100		



## 2.2 학습 데이터들의 패턴 분포 특성

앞에서 분석된 학습 데이터의 패턴 분포 특성을 각 범주별로 종합적으로 분석하면 다음과 같다.

먼저, <표 6>은 밴드 1부터 7까지의 수역에 대한 패턴 특성을 나타낸다. 여기서, 다른 범주도 마찬가지로이지만 밴드 6은 열 밴드이므로 분광 강도가 다른 밴드에 비해 현저히 크게 나타남을 알 수 있다. 또한, 밴드 2와 7은 분광 강도의 패턴이 거의 유사함을 알 수 있다.

<표 6> 수역의 패턴 특성

구 분	밴드 1	밴드 2	밴드 3	밴드 4	밴드 5	밴드 6	밴드 7
최소 분광 강도	103	44	44	29	12	136	5
최대 분광 강도	120	50	76	33	30	161	9
평 균	109.4	46.9	48.9	30.1	16.4	138.5	7.0
표준편차	1.38	1.41	1.12	0.56	0.38	0.48	0.94
중간값	109	47	48	30	17	138	7

<표 7>은 곡물의 패턴 특성을 나타낸다. 곡물의 경우, 밴드 2와 4가 거의 유사한 분포를 나타내고 있다. <표 6>과 <표 7>을 비교했을 때, 밴드 1에서 수역과 곡물의 분광 강도는 거의 유사한 패턴 분포를 나타내고 있는데 이러한 특성은 단일 밴드로 데이터를 분류하지 못하고 대역 조합을 통해 분류를 수행해야 하는 이유가 된다고 앞서 언급했다.

<표 7> 곡물의 패턴 특성

구 분	밴드 1	밴드 2	밴드 3	밴드 4	밴드 5	밴드 6	밴드 7
최소 분광 강도	103	47	48	46	21	156	8
최대 분광 강도	128	52	88	60	156	176	27
평 균	108.1	49.4	57.2	52.3	58.6	161.2	14.4
표준편차	0.67	0.99	1.78	2.80	5.31	1.77	4.28
중간값	108	49.5	57	52	60	161	14

<표 8>은 숲의 패턴 특성을 나타낸다.

〈표 8〉 숲의 패턴 특성

구 분	밴드 1	밴드 2	밴드 3	밴드 4	밴드 5	밴드 6	밴드 7
최소 분광 강도	101	56	43	61	21	147	47
최대 분광 강도	134	72	101	95	194	182	122
평 균	109.6	64.5	65.9	82.7	106.6	167.3	91.3
표준편차	6.66	3.06	9.9	6.57	12.4	8.41	14.44
중간값	109	64	66	83	107	167	94

〈표 9〉는 도시 지역의 패턴 특성을 나타낸다. 도시 지역의 경우 밴드 2, 4, 5, 7은 거의 완벽한 정규 분포를 나타내고 있다.

〈표 9〉 도시 지역의 패턴 특성

구 분	밴드 1	밴드 2	밴드 3	밴드 4	밴드 5	밴드 6	밴드 7
최소 분광 강도	99	41	39	74	19	143	21
최대 분광 강도	126	53	103	111	198	177	53
평 균	108.4	45.3	53.2	90.8	83.2	161.3	31.0
표준편차	3.13	3.09	5.14	8.66	6.78	3.47	9.51
중간값	108	44	53	90	83	161	27

### 3. 실험 데이터들의 분포 특성

앞의 2절에서는 학습 데이터들의 분포 특성을 표로 나타냈다. 본 절에서는 실험에 사용될  $136 \times 136 = 18,496$ 개의 화소들에 대한 분포 특성을 나타낸다. 그러나, 18,000여 개의 화소들에 대한 히스토그램을 모두 나타내기 어려워서 본 논문에서는 패턴 분포 특성을 요약하여 표로 나타내기로 한다.

〈표 10〉은 수역의 분포 특성표이다. 수역의 분포 특성은 전체적으로 평균점을 중심으로 날카롭게 분포되어 있음을 알 수 있다. 물론, 최소 분광 강도와 최대 분광 강도간의 범위가 매우 넓은 밴드도 있긴 하나 분광 강도의 값이 큰 쪽에서는 화소들이 매우 간헐적으로 존재하기 때문에 전반적으로 분광 강도 100 이하에서 정규 분포를 형성하고 있다.

〈표 10〉 실험 데이터 중 수역의 패턴 특성

구 분	밴드 1	밴드 2	밴드 3	밴드 4	밴드 5	밴드 6	밴드 7
최소 분광 강도	99	41	41	25	12	136	3
최대 분광 강도	131	67	89	116	144	170	92
평 균	108.1	45.7	48.2	30.3	16.1	137.5	7.2
표준편차	3.11	3.03	3.13	8.99	9.98	2.34	6.99
중간값	109	46	48	30	16	138	7

〈표 11〉은 곡물의 분포 특성표이다. 수역과는 달리 전체적으로 분광 강도가 낮은 방향으로 편향되어 있는 분포를 나타내고 있다.

〈표 11〉 실험 데이터 중 곡물의 패턴 특성

구 분	밴드 1	밴드 2	밴드 3	밴드 4	밴드 5	밴드 6	밴드 7
최소 분광 강도	99	43	40	42	18	149	6
최대 분광 강도	153	79	106	127	194	183	160
평 균	108.4	49.3	55.1	49.9	27.6	158.3	11.0
표준편차	3.13	3.20	5.04	9.16	12.67	3.11	10.44
중간값	108	49	55	50	27	158	11

〈표 12〉는 숲의 분포 특성표이다. 실험 지역이 야산과 숲이 많은 지역이었기 때문에 다른 범주에 비해 화소의 수가 많은 것이 특징이다. 그러나, 역시 각 밴드별로 정점을 중심으로 거의 정규분포를 형성하고 있다는 특징이 있다.

〈표 12〉 실험 데이터 중 숲의 패턴 특성

구 분	밴드 1	밴드 2	밴드 3	밴드 4	밴드 5	밴드 6	밴드 7
최소 분광 강도	97	40	39	29	17	141	7
최대 분광 강도	134	71	103	126	196	181	118
평 균	101.4	43.9	43.1	84.0	83.2	153.6	23.9
표준편차	2.33	3.09	4.14	8.66	11.78	3.47	9.51
중간값	101	44	43	84	83	154	24

〈표 13〉은 도시 지역의 분포 특성표이다. 도시 지역 역시 거의 정규분포를 이루고 있으나 대역 6만이 분광 강도 163과 분광 강도 173 근처에서 정점을 이루는 쌍봉형인 특

정이 있다.

〈표 13〉 실험 데이터 중 도시 지역의 패턴 특성

구 분	밴드 1	밴드 2	밴드 3	밴드 4	밴드 5	밴드 6	밴드 7
최소 분광 강도	99	42	41	26	17	144	8
최대 분광 강도	146	77	106	124	206	184	182
평 균	116.4	53.2	63.2	83.3	97.1	173.5	47.2
표준편차	3.13	3.09	5.09	8.66	16.31	3.33	9.51
중간값	116	54	64	83	97	173	47

#### 4. Rough 집합이론을 이용한 밴드 분류

이와 같은 분포 특성을 갖는 데이터를 3장의 밴드 분류 알고리즘에 의해 시뮬레이션 프로그램으로 실험을 수행했다. 밴드들에 대한 동치 클래스를 생성하기 위해 분류 기준이 되는 화소의 수는 1개부터 10개까지로 정했다. 이때, 분류 기준이 되는 화소의 선정은 임의로 했다. 그리고, 동일한 조건에서의 실험 횟수는 모두 1,000회를 반복했다. 이와 같은 조건에서 수행한 실험 결과를 요약하면 다음과 같다. 대역 6은 거의 공통적으로 식별 가능, 즉 조합 불가능한 밴드로 판명되었다. 다른 밴드에 비해 분광 강도의 값이 전체적으로 높게 나타났기 때문이다. 이는 LANDSAT TM의 경우 대역 6은 열 밴드(thermal band)이기 때문인 것으로 사료된다. 또한, 분류 기준이 되는 화소의 수가 과다하지 않다면(10개 이하) 전체적으로 밴드 2, 4, 7은 식별 불가능한 밴드로 분석되어 조합 가능하다는 결론을 내릴 수 있었다. 물론 분류 기준이 되는 화소의 수를 줄이면 다른 밴드도 조합 가능할 수 있었으나 그 수가 늘어날수록 점차 밴드 2, 4, 7로 모이는 경향이 발견되었다. 단, 분류 기준이 되는 화소의 수가 10개를 초과할 정도로 과다하면 그 어느 밴드 식별 가능한 결과를 생성하여 군집화 경향을 찾기 어려움을 알 수 있었고 따라서, 밴드 선택을 위해서는 분류 기준이 되는 화소의 수를 10개 이하로 정해야 함을 알 수 있었다.

〈표 14〉는 분류 기준이 되는 화소의 수를 3~5개로 했을 때 나타나는 군집화 경향이 다. 화소의 수가 그리 많지 않기 때문에 다양한 군집화 경향이 나타나는데 그 중 밴드 2, 4, 7로 군집화되는 경향이 뚜렷함을 알 수 있다.

〈표 14〉 실험 결과(화소 수 : 3~5개)

화소 수 3개	군집화 경향	[B <sub>1</sub> , B <sub>3</sub> ]	[B <sub>2</sub> , B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>5</sub> ]	[B <sub>1</sub> , B <sub>7</sub> ]	기타	계
	빈도	101	614	119	53	113	1,000
	비중	10.1%	61.4%	11.9%	5.3%	11.3%	100%
화소 수 4개	군집화 경향	[B <sub>1</sub> , B <sub>2</sub> ]	[B <sub>1</sub> , B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>2</sub> , B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>2</sub> , B <sub>5</sub> ]	기타	계
	빈도	99	214	549	111	27	1,000
	비중	9.9%	21.4%	54.9%	11.1%	2.7%	100%
화소 수 5개	군집화 경향	[B <sub>1</sub> , B <sub>3</sub> ]	[B <sub>1</sub> , B <sub>5</sub> ]	[B <sub>2</sub> , B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>3</sub> , B <sub>5</sub> ]	기타	계
	빈도	136	152	492	103	117	1,000
	비중	13.6%	15.2%	49.2%	10.3%	11.7%	100%

〈표 15〉는 다시 분류 기준이 되는 화소의 수를 6~7개로 증가시킨 것이다. 이제는 3개의 밴드가 군집화되는 경향을 찾아보기 힘들다. 이것은 분류 기준이 되는 화소의 수가 3개 이상의 밴드를 군집화하기에는 어려울 만큼 증가했음을 의미한다.

〈표 15〉 실험 결과(화소 수 : 6~7개)

화소 수 6개	군집화 경향	[B <sub>1</sub> , B <sub>3</sub> ]	[B <sub>1</sub> , B <sub>5</sub> ]	[B <sub>2</sub> , B <sub>7</sub> ]	[B <sub>2</sub> , B <sub>4</sub> ]	기타	계
	빈도	89	113	97	126	575	1,000
	비중	8.9%	11.3%	9.7%	12.6%	57.5%	100%
화소 수 7개	군집화 경향	[B <sub>1</sub> , B <sub>2</sub> ]	[B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>5</sub> , B <sub>7</sub> ]	[B <sub>2</sub> , B <sub>4</sub> ]	기타	계
	빈도	76	99	65	89	671	1,000
	비중	7.6%	9.9%	6.5%	8.9%	67.1%	100%

〈표 16〉은 분류 기준이 되는 화소의 수를 8~10개로 한 경우이다. 전체적으로 10% 이상의 비중을 차지하는 군집화 경향이 발생하지 않기 때문에 분류 기준이 되는 화소의 수를 8개 이상으로 설정한 경우는 군집화 정도를 측정하기 매우 어려운 것으로 판단된다.

〈표 16〉 실험 결과(화소 수 : 8~10개)

화소 수 8개	군집화 경향	[B <sub>2</sub> , B <sub>4</sub> ]	[B <sub>2</sub> , B <sub>7</sub> ]	[B <sub>1</sub> , B <sub>3</sub> ]	기타	계
	빈도	72	91	65	772	1,000
	비중	7.2%	9.1%	6.5%	77.2%	100%
화소 수 9개	군집화 경향	[B <sub>2</sub> , B <sub>4</sub> ]	[B <sub>2</sub> , B <sub>7</sub> ]	[B <sub>3</sub> , B <sub>5</sub> ]	기타	계
	빈도	79	88	53	780	1,000
	비중	7.9%	8.8%	5.3%	78.0%	100%
화소 수 10개	군집화 경향	[B <sub>2</sub> , B <sub>4</sub> ]	[B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>1</sub> , B <sub>3</sub> ]	기타	계
	빈도	66	58	43	833	1,000
	비중	6.6%	5.8%	4.3%	83.3%	100%

〈표 17〉은 지금까지의 실험 결과를 종합해본 것이다. 이를 요약하면 분류 기준이 되는 화소의 수가 2개 이하인 경우는 너무나 다양한 형태로 군집화되기 때문에 여기서 군집화 경향을 찾기 어렵다. 또한, 분류 기준이 되는 화소의 수가 10개를 넘게 되면 모든 밴드들이 상호 식별 가능해지기 때문에 어느 밴드와도 군집화되지 않게 된다. 그러나, 분류 기준이 되는 화소의 수가 3~5개인 경우는 밴드 2, 4, 7간에 군집화되는 경향이 뚜렷하고 또한, 밴드 1, 3으로 군집화되는 경향도 빈번히 발생하고 있다. 화소의 수를 6~8개로 증가시키면 3개의 밴드가 군집화되는 경우는 발생하지 않고 전반적으로 밴드 1과 2를 중심으로 군집화되고 있고, 화소의 수를 9~10개로 증가시키면 다시 밴드 2, 4 중심으로 군집화되는 것을 발견할 수 있다.

〈표 17〉 실험의 종합적 결과

화소 수	2개 이하	3~5개	6~8개	8~10개	10개 초과
군집화 경향	불규칙한 군집화	[B <sub>1</sub> , B <sub>3</sub> ] [B <sub>2</sub> , B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>1</sub> , B <sub>5</sub> ] [B <sub>1</sub> , B <sub>3</sub> ] [B <sub>2</sub> , B <sub>7</sub> ] [B <sub>2</sub> , B <sub>4</sub> ]	[B <sub>2</sub> , B <sub>4</sub> ]	불규칙한 군집화

## 5. 특수한 분포를 갖는 데이터들의 밴드 분류

본 절에서는 앞 절에서 문제로 제기되었던 넓은 분광 강도에 고르게 분포되는 경우와 두 개의 정점을 갖는 쌍봉형 분포와 같은 특수한 특성을 나타내는 분포에 대해 기술한다.

## 5.1 넓은 분광 강도에 고르게 분포된 경우

넓은 분광 강도에 고르게 화소가 분포되는 경우에 해당되는 경우는 밴드 2에서 숲이 보인 분포 특성과 같은 경우이다. 이 경우 분광 강도 44에서 최대의 빈도를 보이며 그 밖의 분광 강도 값에서는 거의 비슷한 빈도를 나타냈다. 그래서 이번 실험에서는 다른 7개의 밴드 중 밴드 2만 추출하고, 숲, 수역, 도시 지역, 곡물의 4개의 범주 중 숲만 추출하여 다시 참조표본을 만들고 이에 대해 동일한 실험을 수행해보았다. 분류 기준이 되는 화소의 수를 앞 절과 같이 다양하게 하지는 않고 3개~8개에 해당되는 경우만 1,000회 반복 수행해보았는데 그 이유는 정규분포 혹은 그와 유사한 분포에서도 화소의 수가 3개~8개인 경우가 가장 균집화 경향이 뚜렷했기 때문이다. 실험 결과 다음과 같은 균집화를 형성하는 것으로 나타났다.

먼저, 분류 기준이 되는 화소의 수가 3~5개인 경우는 다음 <표 18>과 같은 분류 결과를 나타내고 있다. 균집화 경향이 <표 14>와는 약간 다르지만 전체적으로 밴드 2, 4, 7을 중심으로 균집화되는 것을 알 수 있다. 특히, 분류 기준이 되는 화소의 수가 5개인 경우는 3개 이상의 밴드가 균집화되는 경향을 발견할 수 없는데 이는 앞의 실험에서 분류 기준이 되는 화소의 수가 6개 이상인 경우부터 발생했던 현상이다. 이는 정규 분포와 유사한 분포를 이루는 분포 특성이 왜곡될수록 균집화 정도가 낮아짐을 의미한다.

<표 18> 실험 결과(화소 수 : 3~5개)

화소 수	균집화 경향	[B <sub>1</sub> , B <sub>2</sub> ]	[B <sub>2</sub> , B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>4</sub> ]	[B <sub>2</sub> , B <sub>7</sub> ]	기타	계
	3개	빈도	89	543	98	42	228
비중		8.9%	54.3%	9.8%	4.2%	22.8%	100%
화소 수 4개	균집화 경향	[B <sub>2</sub> , B <sub>7</sub> ]	[B <sub>1</sub> , B <sub>5</sub> , B <sub>7</sub> ]	[B <sub>2</sub> , B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>2</sub> , B <sub>5</sub> ]	기타	계
	빈도	47	103	491	56	303	1,000
	비중	4.7%	10.3%	49.1%	5.6%	30.3%	100%
화소 수 5개	균집화 경향	[B <sub>1</sub> , B <sub>5</sub> ]	[B <sub>1</sub> , B <sub>7</sub> ]	[B <sub>2</sub> , B <sub>7</sub> ]	[B <sub>3</sub> , B <sub>4</sub> ]	기타	계
	빈도	58	100	223	87	532	1,000
	비중	5.8%	10.0%	22.3%	8.7%	53.2%	100%

분류 기준이 되는 화소의 수를 6~8개로 한 경우는 다음 <표 19>로 알 수 있다. <표 15>의 결과와 비교했을 때, 균집화 경향은 비슷하나 밴드 2와 5의 균집화 경향이 현저히 떨어져 기타로 분류되어 있음을 알 수 있다. 특히, 분류 기준이 되는 화소의 수가 8

개로 늘어난 경우 오직 한 종류의 군집화 경향만 발생하는데 이는 <표 15>의 실험 결과와 비교했을 때, 군집력이 현저히 떨어졌음을 알 수 있다. 따라서, 분류 기준이 되는 화소의 수를 증가시킬 수록 특정 군집화 경향이 나타나기 어려움을 알 수 있었다.

<표 19> 실험 결과(화소 수 : 6~8개)

화소 수 6개	군집화 경향	[B <sub>1</sub> , B <sub>3</sub> ]	[B <sub>2</sub> , B <sub>7</sub> ]	[B <sub>2</sub> , B <sub>4</sub> ]	기타	계
	빈도	55	73	99	773	1,000
	비중	5.5%	7.3%	9.9%	77.3%	100%
화소 수 7개	군집화 경향	[B <sub>1</sub> , B <sub>4</sub> ]	[B <sub>2</sub> , B <sub>4</sub> ]		기타	계
	빈도	44	56		900	1,000
	비중	4.4%	5.6%		90.0%	100%
화소 수 8개	군집화 경향	[B <sub>1</sub> , B <sub>3</sub> ]			기타	계
	빈도	35			965	1,000
	비중	3.5%			96.5%	100%

본 실험에서는 결론적으로 다음과 같은 군집화 경향으로 요약할 수 있다. 즉, 분류 기준이 되는 화소의 수와 관계없이 전반적으로 밴드 2, 4, 7 중심으로 군집화되는 경향이 있으나 앞 절의 실험에서 사용한 7개의 밴드와 4개의 범주를 이용한 경우보다는 군집력이 떨어짐을 알 수 있다. 이는 밴드 2에서 숲의 분포 특성이 고르게 분포되는 특성이기 때문에 평균과 표준편차의 중심을 이탈하는 데이터들이 앞의 실험보다 많기 때문으로 분석된다. 그러나, 밴드 2, 4, 7 중심의 군집화 경향이 여전히 나타나는 것으로 보아 이와 같은 분포 특성을 갖는 경우도 본 논문에서 제안하는 방법이 유용한 것으로 나타났다.

<표 20> 실험의 종합적 결과

화소 수	3~5개	6~8개
군집화 경향	[B <sub>2</sub> , B <sub>7</sub> ] [B <sub>2</sub> , B <sub>4</sub> , B <sub>7</sub> ]	[B <sub>1</sub> , B <sub>3</sub> ] [B <sub>2</sub> , B <sub>4</sub> ]

## 5.2 쌍봉형 분포 특성을 갖는 경우

쌍봉형 분포를 갖는 경우는 밴드 4에서 곡물에 해당되는 학습 데이터들이다. 이는 분



광 강도 14를 중심으로 정규 분포를 이루고, 또 분광 강도 52를 중심으로 또 하나의 정규 분포를 형성하는 전형적인 쌍봉형 분포 구조이다. 만일, 이들 중 한쪽의 정규 분포가 다른 한쪽의 정규 분포에 비해 현저히 왜소하다면 이는 본 논문에서 제안하는 방법을 그대로 적용할 수 있을 것으로 판단된다. 그러나, 본 실험의 경우는 거의 동등한 비중을 차지하는 대칭적 분포 구조이다. 그래서, 본 논문에서는 밴드 4와 곡물이라는 범주만 추출하여 이들의 대표값을 최빈값으로 하여 실험을 수행했다. 그러나, 최빈값으로 구성된 참조표를 통해 동치 클래스를 생성하여 군집화 경향을 조사한 결과 특정 밴드들이 군집화하지 않는 것을 발견했다. 그래서, 다시 본 실험에서 사용한 평균값과 표준편차를 이용하여 이들에 대해서만 실험을 수행한 결과 다음과 같은 결과가 생성되었다. 단, 전체적인 군집화 경향만을 확인하기 위해 분류 기준이 되는 화소의 수를 3~5개로 축소하였고 실험 횟수도 100회로 축소하여 실험을 수행했다.

분류 기준이 되는 화소의 수가 3개인 경우는 다음과 같은 실험 결과를 보였다.

〈표 21〉 실험 결과(화소 수 : 3개)

군집화 경향	[B <sub>2</sub> , B <sub>4</sub> ]	[B <sub>1</sub> , B <sub>4</sub> ]	[B <sub>2</sub> , B <sub>7</sub> ]	기타	계
빈 도	27	19	9	45	100
비 중	27.0%	19.0%	9.0%	45.0%	100%

분류 기준이 되는 화소의 수가 4개인 경우는 역시 다음과 같은 결과를 생성했다.

〈표 22〉 실험 결과(화소 수 : 4개)

군집화 경향	[B <sub>2</sub> , B <sub>7</sub> ]	[B <sub>2</sub> , B <sub>4</sub> ]	기타	계
빈 도	14	8	78	100
비 중	14.0%	8.0%	78.0%	100%

마지막으로, 분류 기준이 되는 화소의 수를 5개로 한 결과는 다음과 같다.

〈표 23〉 실험 결과(화소 수 : 5개)

군집화 경향	[B <sub>1</sub> , B <sub>4</sub> ]	[B <sub>2</sub> , B <sub>4</sub> ]	[B <sub>1</sub> , B <sub>7</sub> ]	기타	계
빈 도	12	8	8	72	100
비 중	12.0%	8.0%	8.0%	72.0%	100%

본 실험 결과를 종합적으로 분석해 보면 평균과 표준편차를 이용한 동치 클래스 생성 방법이 결과의 빈도는 매우 떨어지나 다른 방법과 비교하여 상대적으로 다양한 분포를 갖는 데이터들을 대표할 수 있는 대표값으로 가장 적절한 것으로 생각되고 또한, 이를 통해 정규 분포나 평범한 분포를 갖는 경우와 유사하게 군집화 경향을 보임을 알 수 있다.

## 6. 실험 결과의 평가

다중 분광 이미지 데이터의 양호한 분석 정확도를 보장하기 위해서는 탐사에 이용된 모든 밴드를 데이터 분석에 사용해야 할 것이다. 그러나, 이 경우 실시간내에 처리할 수 없을 정도의 높은 연산 오버헤드가 발생하기 때문에 일반적으로 ERDAS와 같은 대부분의 원격 탐사 데이터 처리 도구들은 수역, 곡물, 숲, 도시 지역 등을 분류하는데 대역 2, 4, 7을 사용한다. 본 논문에서는 Rough 집합을 이용하여 밴드의 특징 추출을 수행한 결과, 분류 기준이 되는 화소의 수에 따라 차이가 발생하긴 했지만 전반적으로 밴드 2, 4, 7은 상호 식별 불능 관계임을 알 수 있었으며 이들은 상호 조합 가능한 관계로 해석할 수 있었다. 이를 통해, Rough 집합을 이용한 결과가 기존의 물리적·파장적 특성을 중심으로 한 밴드 선택 방법과 거의 동일하게 군집화되는 경향을 확인할 수 있었다. 그리고 분류 기준이 되는 화소의 수가 2~10개인 제한된 문제를 대상으로 실험해본 결과 분류 기준이 되는 화소의 수가 5개 이하인 경우는 주로 밴드 2, 4, 7을 중심으로 군집화되는 경향을 보였고, 6개 이상 10개 이하인 경우에는 3개의 밴드가 동시에 군집화되는 경우는 없었으나 밴드 2, 4, 7 중의 두 밴드가 조합되는 것을 확인하였다. 이상의 결과로부터 Rough 집합을 이용하면 이 집합의 불충분하거나 일관성이 없는 데이터에 대한 분류를 용이하게 한다는 특징 때문에 밴드의 수가 매우 많은 초다중 분광 밴드에서도 유효 밴드 분류가 용이할 것으로 생각된다.

## IV. 결 론

본 논문에서는 다중 분광 환경의 유효 밴드 분류를 위해 Rough 집합이론을 도입한 새로운 밴드 분류 알고리즘을 제안했다. 제안된 방법은 Rough 집합이론의 데이터 분류

및 식별 능력을 이용하였고, 이 방법을 LANDSAT TM으로 수집한 1992년 6월 2일 한강의 행주대교 인근 지역의 데이터를 대상으로 시뮬레이션하여 그 타당성 및 성능을 확인하였다. 시뮬레이션 결과, 기존의 많은 시행착오를 거쳐 얻은 밴드 분류 결과와 거의 같은 밴드를 자동적으로 추출함으로써 조작자의 큰 노력없이도 최적의 밴드를 선택할 수 있었고, 보다 효과적인 데이터 분석에 활용할 수 있음을 확인하였다.

제안된 분류 알고리즘은 비록 다중 분광 환경하에서 수집한 데이터를 이용하여 실험했지만 초다중 분광 환경으로 상황이 확장되었을 때도 그대로 활용함으로써 좀 더 용이한 분류를 통해 이미지 데이터의 분석 및 분류 효율을 높일 수 있을 것으로 사료된다. 아직까지는 7개의 밴드를 갖는 LANDSAT TM 데이터나 4개의 밴드를 갖는 SPOT HRV 데이터를 이용하여 주로 처리하고 이들 결과를 GIS의 입력 데이터로 활용하고 있지만 제안된 알고리즘을 이용하여 초다중 분광 데이터를 처리하여 GIS를 구축하면 아주 높은 분석 정확도를 갖는 강력한 GIS를 구축할 수 있을 것이다.

향후로는, 밴드 분류뿐 아니라 각 화소의 분류 과정에도 Rough 집합이론을 도입하여 분류 정확도를 향상시킬 수 있는 방법이 모색될 수 있는 연구가 진행되어야 할 것이다.

## 참 고 문 헌

1. S. Abe and R. Thawonmas, "A Fuzzy Classifier with Ellipsoidal Regions," *IEEE Trans. on Fuzzy Systems*, Vol. 5, No. 3, pp. 358~368, Aug., 1997.
2. H. Bischof, W. Schneider and A.J. Pinz, "Multispectral Classification of LANDSAT-Images using Neural Networks," *IEEE Trans. on Geosci & Remote Sensing*, Vol. 30, No. 3, pp. 482~490, May, 1992.
3. F. Canters, "Evaluating the Uncertainty of Area Estimates Derived from Fuzzy Land-Cover Classification," *Photogrammetric Engineering & Remote Sensing*, Vol. 63, No. 4, pp. 403~414, Apr., 1997.
4. R.B. Erb, "ERTS-1 Urban Land Use Analysis," *Report for Period-June 1973, in the ERTS-1 Investigation(ER-600)*, Vol. V, NASA TM X-58121.
5. P.D. Heermann and N. Khazenie, "Classification of Multispectral Remote Sensing Data Using a Back-Propagation Neural Network," *IEEE Trans. on Geosci & Remote Sensing*, Vol. 30, No. 1, pp. 81~88, Jan., 1992.
6. A.R. Huete, "Separation of Soil-Plant Spectral Mixtures by Factor Analysis," *Remote Sensing of Environmental*, Vol. 19, pp. 237~251, 1986.
7. L. Jimenez and D. Landgrebe, "Supervised Classification in High Dimensional Space : Geometrical, Statistical and Asymptotical Properties of Multivariate Data," *IEEE Trans. on Systems, Man, and Cybernetics*, Jan., 1998.
8. J.I. Kim and S.C. Kim, "A Fuzzy Partitioning Method of Spectral Space for Remote Sensing Image Classification," *Proceedings of FUZZ-IEEE/IFES '95 Joint Conference*, Vol. 3, pp. 1,125~1,130, 1995.
9. F.J. Knorr and J.H. Futrell, "Separation of Mass Spectra of Mixtures by Factor Analysis," *Anal. Chem.*, Vol. 51, pp. 1236~1241.
10. T.M. Lillesand and R.W. Kiefer, *Remote Sensing and Image Interpretation*, 2nd Ed., John Wiley & Sons, 1987.
11. P.M. Mather, *Computer Processing of Remotely-Sensed Image*, John Wiley & Sons Inc., 1993.
12. M. Nishida, K. Hayashi, N. Yoshimura and K. Ohtsuka, "Application of Fuzzy Reasoning to Estimation of Class Mixture Proportion of Mixed Pixel on

- Remote Sensing," *IGARSS '93*, Vol. 4, pp. 1,965~1,968, Tokyo, Japan, Aug. 18-21, 1993.
13. K. Okamodo and M. Fukuhara, "Improvement in the Accuracy of Paddy Field Area Estimation using the Area Ratio of Each Pixel Derived from LANDSAT TM Data," *IGARSS '93*, Vol. 4, pp. 1,699~1,700, Tokyo, Japan, Aug., pp. 18~21, 1993.
14. Z. Pawlak, *Rough Sets, International Journal of Computer and Information Science*, 1982.
15. Z. Pawlak, *Rough Sets, Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, 1991.
16. P. Rosenberg, "Resolution, Detectability and Recognizability," *PE & RS*, Vol. 37, No. 12, pp. 1,255~1,258, 1971.
17. R.A. Schowengerdt, *Techniques for Image Processing and Classification in Remote Sensing*, Academic Press, 1983.
18. H. Shimoda, Y. Yamaguchi, K. Fukue and T. Sakata, "Effects of Spatial Resolution to Landcover Classification Accuracies for SPOT HRV and LANDSAT TM Data," *ISPRS Kyoto '88*, Vol. 27, Part B7, Commission VII, pp. 64~76, 1988.
19. S.H. Won, J.I. Kim and S.C. Kim, "A Classification Algorithm of Mixed Pixel using Fuzzy Number in Geographic Information System," *Proceedings of JTC/CSCC '94*, Vol. 1, pp. 244~247, Jul., 1994.
20. 김진일, "위성 탐사 이미지 분류를 위한 다중 분광 패턴 공간에서의 퍼지 규칙 생성에 관한 연구," 서강대학교 대학원, 박사학위논문, 1994.
21. 양영규 외 9인, "원격 탐사 관련 기술 실용화 연구( I, II, III )," 한국과학기술연구원, 시스템공학센터, 1989~1991.
22. 원성현 · 이병성 · 정환묵, "러프 집합을 이용한 다중 분광 이미지 데이터의 분류," 한국퍼지및지능시스템학회 학술발표논문집, 제7권, 제1호, pp. 205~208, 1997.
23. 원성현 · 정환묵, "Rough 집합이론을 이용한 원격 탐사 다중 분광 이미지 데이터의 특징 추출," 한국퍼지및지능시스템학회 논문지, 제8권, 제3호, pp. 16~25, 1998.

## Abstract

### A Study on Classifications of Remote Sensed Multispectral Image Data using Soft Computing Technique - Stressed on Rough Sets

Won, Sung-hyun

Processing techniques of remote sensed image data using computer have been recognized very necessary techniques to all social fields, such as, environmental observation, land cultivation, resource investigation, military trend grasp and agricultural product estimation, etc. Especially, accurate classification and analysis to remote sensed image data are important elements that can determine reliability of remote sensed image data processing systems, and many researches have been processed to improve these accuracy of classification and analysis. Traditionally, remote sensed image data processing systems have been processed 2 or 3 selected bands in multiple bands, in this time, their selection criterions are statistical separability or wavelength properties. But, it have be bring up the necessity of bands selection method by data distribution characteristics than traditional bands selection by wavelength properties or statistical separability. Because data sensing environments change from multispectral environments to hyperspectral environments.

In this paper, for efficient data classification in multispectral bands environment, a band feature extraction method using the Rough sets theory is proposed. First, we make a look up table from training data, and analyze the properties of experimental multispectral image data, then select the efficient band using indiscernibility relation of Rough set theory from analysis results. Proposed method is applied to LANDSAT TM data on 2 June 1992. From this, we show clustering trends that similar to traditional band selection results by wavelength properties, from this, we verify that can use the proposed method that centered

on data properties to select the efficient bands, though data sensing environment change to hyperspectral band environments.