

## 지

구 혹은 지구 밖의 물체에 대한 원격탐사 자료를 분석하여 주제 정보를 추출할 수 있는데, 이때 자료가 정보로 변환된다. 원격탐사는 국지적 및 전지구적 규모의 토지이용과 토지피복 변화를 파악하기 위한 정보의 원천으로서 많이 사용되어 왔다(Townshend and Justice, 2002 ; Lunetta and Lyons, 2003). 통계적 패턴인식 기법에 기초한 다중분광 원격탐사 영상의 토지이용/토지피복 분류가 다양한 정보 추출 기법들 중 가장 흔히 쓰이는 방법이다(Narumalani et al., 2002). 이 방법은 대상 지역의 영상이 전자기 스펙트럼의 다중밴드에서 수집되고 기하보정이 잘 이루어졌다는 가정 하에 이루어진다. 디지털 다중분광 원격탐사 영상에서 토지피복 정보를 추출하기 위한 일반적인 단계가 그림 9-1에 정리되어 있다.

실제로 다중분광 영상의 분류는 아래와 같은 다양한 방법들을 사용하여 수행될 수 있다(그림 9-1과 9-2).

- 매개변수 및 비매개변수를 이용한 통계값에 기초한 알고리즘(비율척도와 등간척도 자료에 사용)과 명목척도 자료에도 사용할 수 있는 비계량 방법에 기초한 알고리즘(Duda et al., 2001)
- 감독, 무감독 분류 논리의 사용
- 범주형 혹은 퍼지 주제정보를 만들기 위한 범주형 혹은 퍼지 분류 논리의 사용
- 화소단위나 객체지향 분류 논리의 사용
- 혼합적인 접근법

최대우도법과 무감독 분류와 같은 매개변수(*parametric*)를 이용한 방법은 원격탐사 자료가 정규분포를 이루고 분류를 하기 위한 자료의 밀도 함수를 알고 있다고 가정한다(Duda et al., 2001). 최근린 분류자, 퍼지 분류자, 신경망과 같은 비매개변수(*nonparametric*)를 이용한 방법은 정규분포를 띠지 않는 원격탐사 자료에 적용될 수 있으며 밀도함수를 알아야 한다는 가정이 필요없다(예, Friedl et al., 2002 ; Liu et al., 2002 ; Qui and and Jensen, 2004). 규칙에 기초한 의사결정분지

도(*decision tree*) 분류자와 같은 비계량(*nonmetric*) 방법은 0~100%의 반사도와 같은 실수값 자료나 클래스 1은 산림, 클래스 2는 농경지 등과 같은 명목척도 자료에도 적용할 수 있다(예, Tullis and Jensen, 2003 ; Stow et al., 2003).

감독분류(*supervised classification*)에서 도심지, 농경지, 습지 등의 토지피복 형태와 그 위치는 현장 확인, 항공사진의 판독, 지도 분석, 개인의 경험 등의 조합을 통하여 선형적으로 알고 있다(Hodgson et al., 2003). 분석가는 원격탐사 자료에 이러한 알려진 토지피복 중 균일한 예들을 대표하는 특정 위치를 찾으려고 한다. 이러한 지역을 보통 훈련지역(*training sites*)이라고 하는데, 이는 이 지역의 분광 특성이 영상의 남은 부분의 토지피복도를 만들기 위한 분류 알고리즘을 훈련하는데 사용되기 때문이다. 평균, 표준편차, 공분산 행렬, 상관 행렬 등과 같은 다변량 통계 변수들을 각각의 훈련 지역에 대하여 계산하여, 이를 기초로 훈련지역 안과 밖의 모든 화소를 평가한 뒤 각 화소를 가장 확률이 높은 클래스에 할당한다.

무감독 분류(*unsupervised classification*)에서는 지상의 참조정보가 부족하거나 영상의 표면 구조물들이 제대로 정의되어 있지 않아서 클래스별 토지 피복 형태를 선형적으로 알 수 없다. 통계적으로 정의된 기준에 따라서 컴퓨터가 비슷한 분광특성을 가진 화소들을 군집(*cluster*)으로 분류한다(Duda et al., 2001). 그 다음에 분석가가 각 군집마다 이름을 붙이고 이러한 분광 군집을 정보 클래스로 통합한다.

감독분류와 무감독 분류 알고리즘은 보통 산림, 농지 등과 같은 이산적인 항목으로 분류된 지도를 만들기 위해 범주형 분류(*hard classification*) 논리를 사용한다(그림 9-2a). 반대로, 실세계의 이질적이고 부정확한 속성을 고려한 퍼지 분류(*fuzzy classification*) 논리도 가능하다(그림 9-2b). 퍼지 분류는 퍼지 정보를 가지고 있는 주제도 결과물을 만들어낸다. 퍼지 분류는 원격탐사 감지

기가 순간시야각(IFOV) 안에서 발견되는 토양, 물, 식물과 같은 생물리적 물질의 이질적 혼합으로부터 반사 혹은 방출된 복사속을 기록한다는 사실을 기반으로 한다(Foody, 1996 ; 2000 ; Karaska et al., 1997). 순간시야각(화소) 내에 발견되는 토지피복 클래스는 종종 뚜렷한 범주형 경계를 가지지 않는다. 즉, 현실은 매우 모호하고 불균일한데, 이를 퍼지하다고 한다(Ji and Jensen, 1999). 퍼지 분류에서는 각 화소가 가능한  $m$ 개의 클래스 중에서 1개의 클래스로 정의되는 대신에  $m$ 개의 멤버십 등급값을 가진다. 멤버십 등급 값은 1개의 화소를 그 안에서 발견되는  $m$ 개의 토지피복 형태의 비율을 나타내는 값으로, 예를 들어 10% 나지, 10% 관목지, 80% 산림으로 나타내는 것을 말한다(그림 9-2b). 이러한 정보는 좀 더 정확한 토지피복 정보를 얻어내는 데 사용될 수 있고, 특히 혼합 화소를 파악할 때 사용된다(Ji and Jensen, 1996 ; Foody, 2002).

과거에는 대부분의 디지털 영상분류가 전체 영상을 화소단위로 처리하는 것에 기초하였는데, 이를 화소단위 분류(*per-pixel classification*)라 부른다(Blaschke and Strobl, 2001). 객체지향 분류(*object oriented classification*) 기법은 다중해상도 영상분할 기법을 이용해 영상을 상대적으로 균등한 영상객체(패치나 조각이라고도 불림)들로 분해하도록 한다(Baatz et al., 2001). 이어 균질한 영상객체의 다양한 통계적 특성들은 전통적인 통계기법이나 퍼지논리 분류에 의해 처리된다. 영상분할(*image segmentation*)에 기초한 객체지향 분류는 1×1m IKONOS 영상과 0.61×0.61m QuickBird 영상과 같이 공간해상도가 높은 영상의 분석에 종종 사용된다.

어떤 패턴 분류 방법도 본질적으로 다른 방법보다 더 뛰어난 것은 없다. 분류시 문제의 본질, 연구대상지역의 생물리적 특성, 원격탐사 자료의 분포(예, 정규분포), 선형적 지식 등을 통해 어떤 분류 알고리즘이 적합한 결과를 도출할 수 있을지 결정한다. Duda et al.(2001)은 “우리는 어떤 특정한 학습 또는 인식 알고리즘의 전반적인

**디지털 원격탐사 자료에서  
주제 토지피복 정보를 추출하는 데 사용되는 일반적 단계**

**토지피복 분류 문제의 속성 제시**

- \* 대상 지역을 구체화
- \* 클래스 정의
- \* 범주형 혹은 퍼지 분류 결정
- \* 화소단위 혹은 객체지향 분류 결정

**적절한 원격탐사 및 초기 지상참조자료 수집**

- \* 다음의 기준에 맞춰 원격탐사 자료 선택
  - 원격탐사 시스템 고려 사항
    - 공간, 분광, 시간, 방사해상도
  - 환경 고려 사항
    - 대기조건, 토양습도, 생물 계절주기 등
- \* 초기 지상참조자료 수집
  - 연구 지역의 선행 지식에 기초

**원격탐사 자료를 처리하여 주제 정보 추출**

- \* 방사보정(혹은 정규화)(제6장)
- \* 기하보정(제7장)
- \* 적절한 영상분류 논리 선택
  - 매개변수(예, 최대우도법, 군집화)
  - 비매개변수(예, 최근린, 신경망)
  - 비계량(예, 규칙기반 의사결정분지도 분류자)
- \* 적절한 영상 분류 알고리즘 선택
  - 감독 분류
    - 평행육면체 분류법, 최소거리 분류법, 최대우도 분류법
    - 기타(하이퍼분광 일치화 필터링, 분광각 매핑-제11장)
  - 무감독 분류
    - 체인방법, 다중패스 ISODATA
    - 기타(퍼지  $c$ -평균)
  - 인공지능을 포함한 혼성분류(제10장)
    - 전문가 시스템 의사결정분지도, 신경망
- \* 초기 훈련지역으로부터 자료 추출(필요시)
- \* 피처선택 기준에 맞춰 최적의 밴드 선택
  - 그래픽(예, 공동분광 도표)
  - 통계(예, 변환발산, TM 거리)
- \* 다음에 기초하여 훈련 통계값 및 규칙 추출
  - 최종 밴드선택(필요시)
  - 기계학습(제10장)
- \* 주제정보 추출
  - 각 화소별 혹은 분할된 각 영상객체별(감독분류)
  - 화소 및 영상객체 명명(무감독 분류)

**정확도 평가 수행(제13장)**

- \* 방법 선택
  - 정성적 신뢰도 구축
  - 통계적 측정
- \* 클래스별 필요한 표본 수 결정
- \* 표집계획 선택
- \* 지상참조 검증 정보 수집
- \* 오차행렬 구축 및 분석
  - 단변량 및 다변량 통계분석

**이전에 설정한 가설을 수용 혹은 기각**

**정확도가 허용가능하면 결과 배포**

그림 9-1 디지털 원격탐사 자료에서 주제 토지피복 정보를 추출하기 위한 일반적인 단계.

범주형 대 퍼지 논리에 기초한 원격탐사 자료 분류

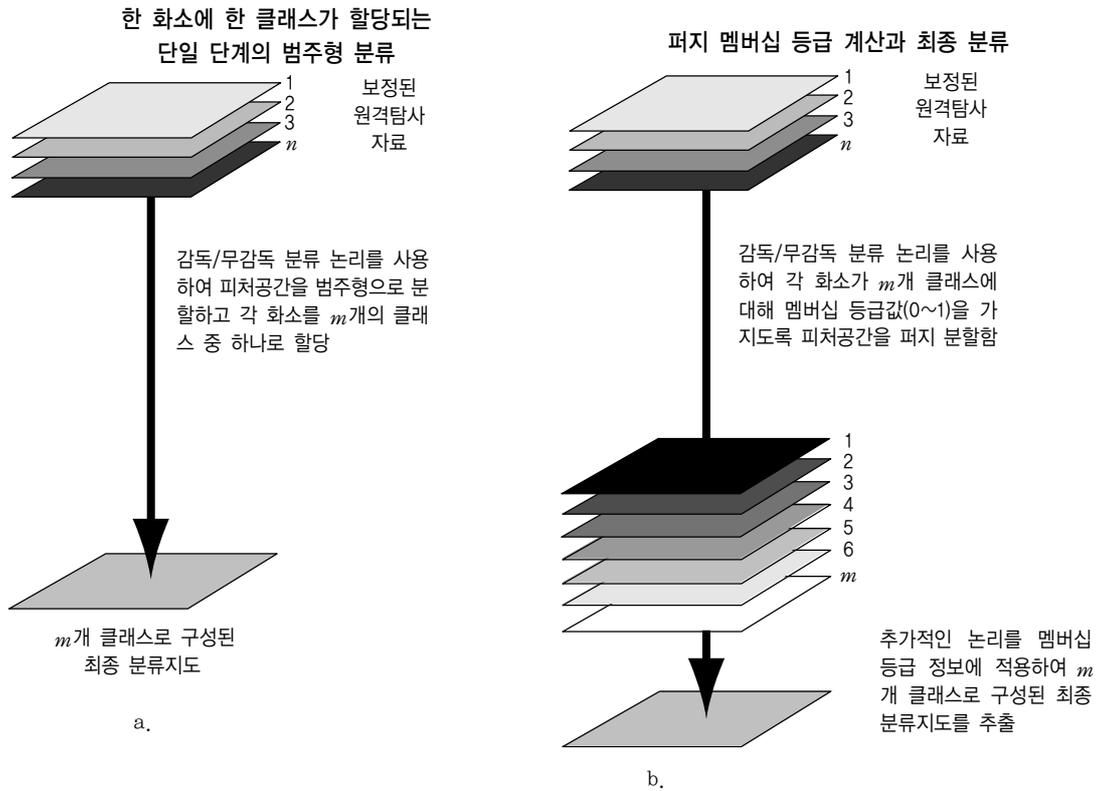
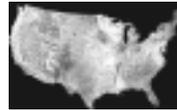


그림 9-2 감독 분류나 무감독 분류를 사용한 전통적인 단일 단계의 범주형 분류와 퍼지 논리를 사용한 분류 사이의 관계.

우수성을 입증하려는 연구에 대해 건전한 의심을 해박야 한다”라고 조언을 하였다.

이 장에서는 매개변수 및 비매개변수 기법을 사용한 감독, 무감독 분류를 포함하여 다중분광 원격탐사 자료로부터 토지피복 정보를 추출하는 데 사용되는 기본적인 통계적인 패턴인식 기법을 소개한다. 퍼지 분류의 기본적인 특성도 소개하며, 끝에는 객체지향 영상분할과 분류 원리를 제시하였다. 신경망(neural network)과 전문가 시스템 의사결정분지도 분류자 등을 포함한 비계량 방법은 제10장에 소개되어 있다(Jensen et al., 2000, 2001). 하이퍼분광 영상으로부터 유용한 정보를 추출하기 위한 알고리즘은 제11장에서 논의된다.



감독 분류

그림 9-1에 나타난 일반적인 단계를 이해하고 적용하면 감독분류 알고리즘을 사용하여 유용한 주제정보를 얻을 수 있다. 분석가는 먼저 가설을 검증하기 위한 대상 지역(Region Of Interest : ROI)을 구체화하여야 한다. 그 다음 분류체계에서 조사할 관심 클래스들을 정의한다.

분석가는 관심 클래스로부터 어떤 결과물(범주형 혹은 퍼지)을 낼 것인지와 어떤 분류 논리(화소단위 혹은 객체지향 분류 논리)를 사용할 것인지를 결정한다 (Green and Congalton, 2003). 그 다음에 분석가는 센서 시스템 및 환경적 제한 요소를 동시에 고려하여 적합한 디지털 원격탐사 자료를 수집한다. 이상적으로 지

상참조정보는 원격탐사 자료의 수집과 동시에 이루어진다. 앞 장에서 언급했듯이, 원격탐사 자료에 방사보정과 기하보정을 수행한다. 이어 적합한 분류 알고리즘을 선택하고, 필요하면 초기 훈련자료를 수집한다. 어떤 훈련 클래스와 다른 클래스를 구별할 수 있는 최적의 다중분광밴드를 결정하기 위한 피처(밴드)를 선택한다. 필요한 경우, 부가적인 훈련자료를 수집하고, 분류 알고리즘을 적용하여 분류지도를 만들어낸다. 엄밀한 정확도 평가(오차 평가)를 수행하여(제13장 참조), 수용할 만한 결과가 나오면 분류지도와 관련 통계들을 관련자와 기관에 배포한다.

### 토지이용 및 토지피복 분류체계

토지피복(*land cover*)이란 물, 모래, 농작물, 산림, 습지, 아스팔트와 같은 인공물처럼 지형에 존재하는 물질의 종류를 말한다. 토지이용(*land use*)이란 농경지, 상업지, 주거지와 같이 토지 표면에 사람들이 무엇을 하느냐를 말한다. 현재 인간에 의한 지구 토지 표면의 변화 속도, 크기, 규모는 인간 역사상 선례가 없을 정도이다. 결론적으로, 토지피복과 토지이용 자료는 산림벌채를 막고 지속가능한 거주지 성장을 유지하고, 수자원의 질과 공급을 보호하는 것 등의 UN의 안건 21(*Agenda 21*)과 같은 이슈에 중추적인 역할을 하는 자료이다(DeFries and Townshend, 1999 ; Jensen et al., 2002). 인간이 경관에 미치는 영향을 고려하여 토지피복과 토지이용의 변화를 평가할 수 있는 기본 데이터셋을 구축할 필요가 있다(Lunetta and Elvidge, 1998).

IGBP(International Geosphere-Biosphere Programme)와 IHDP(International Human Dimensions of Global Environmental Change Programme)에서는 다음과 같은 사항을 제안하였다.

앞으로 수십 년간 토지이용/토지피복 변화가 지구에 끼치

는 영향은 잠재적인 기후변화와 관련된 영향만큼 혹은 그보다 더 중요할지도 모른다. 기후변화 자체와는 다르게 토지이용/토지피복 변화는 지구 환경 변화에 알려져 있고 명백한 부분이다. 이러한 변화에 따른 영향은 현재 잠재적인 기후 온난화에서부터 지반 침식과 생태계의 파괴까지 그리고 식량생산부터 전염병의 확산까지 이른다.

토지피복 자료는 특히 측량될 수 없는 넓은 지역의 개개의 종과 종의 집단의 분포를 예측하는 데 유용하다. 토지피복 자료의 활용성과 정확도가 향상됨에 따라 다양한 예측 모델이 널리 사용되고 있다. 예를 들어, 원격탐사로 얻은 토지피복 정보는 가장 큰 종 분포 모델링인 Gap 분석 프로그램(Gap Analysis Program : GAP)에 사용되고 있는데, 그 목적은 대상 종이 선호하는 서식지의 자세한 지도를 만들고, 식물 계절주기(*plant phenology*)를 모니터링 하는 데 있다(Kerr and Ostrovsky, 2003).

원격탐사 자료를 성공적으로 토지이용 및 토지피복 정보로 분류하기 위해서는 모든 관심 클래스가 주의 깊게 선택되고 정의되어야 한다(Lunetta et al., 1991 ; Congalton and Green, 1999). 이를 위해서 분류체계(*classification scheme*)를 사용하는데 분류체계는 논리적인 기준에 따라 구성된 정보 클래스의 분류학상의 정확한 정의를 담고 있어야 한다. 만약 범주형 분류를 수행하려고 하면 분류 시스템상의 클래스들은 아래의 조건을 만족해야 한다.

- 상호 배타적
- 포괄적
- 계층적

상호 배타적(*mutually exclusive*)이라 함은 분류학적으로 클래스 사이에 중복 혹은 퍼지성이 없다는 것을 의미한다. 예를 들어, 낙엽수와 상록수는 별개의 클래스이므로 중복되지 않는다. 포괄적(*exhaustive*)이라 함

은 모든 토지피복 클래스가 고려되고 어떠한 것도 빠져서는 안 된다는 말이다. 계층적(*hierarchical*)이라 함은 낮은 레벨의 클래스(예, 단독주택 주거지, 다세대주택 주거지)가 주거지라는 더 높은 레벨의 항목에 계층적으로 포함되어야 함을 말한다. 이로부터 단순화된 주제도가 필요한 경우 쉽게 만들 수 있다.

분석가는 정보 클래스와 분광 클래스 간의 기본적인 차이점을 인식해야 한다. 정보 클래스(*information class*)는 인간이 정의하는 것인 반면, 분광 클래스(*spectral class*)는 원격탐사 자료에 내재한 것이고 분석가에 의해 인식되고 명명된다. 예를 들어, 도시지역의 원격탐사 영상에서 단독주택 주거지 형태가 있다고 하자. SPOT (20×20m)과 같이 공간해상도가 낮은 원격탐사 영상은 비교적 순수한 식생화소와 콘크리트 및 아스팔트 도로 화소를 일부 기록할 수 있을 것이다. 그러나 이런 주거 지역에서 화소의 밝기값은 식생과 아스팔트/콘크리트 등이 혼합된 것으로부터의 반사도 함수일 것이다. 불투수층 지도에 관심이 있을 경우를 제외하고는 (1) 콘크리트, (2) 아스팔트, (3) 식생, (4) 식생과 아스팔트/콘크리트의 혼합, 이런 식의 클래스로 나누어진 지도를 원하는 계획가나 행정 전문가가 없다. 오히려 혼합 클래스를 단독주택 주거지라고 재 명명하는 것을 더 선호할 것이다. 반면 분석가는 혼합 클래스와 단독주택 주거지 사이에 관련이 있을 경우만 이렇게 해야 한다. 분석가는 종종 관련 기간의 요구에 따르기 위하여 분광 클래스를 정보 클래스로 변환해야 한다. 분석가는 센서 시스템에 대한 공간특성과 분광특성을 잘 이해하여야 하며 이러한 시스템 매개변수를 영상이나 화소 순간시야각 내에서 발견되는 물질의 종류 및 비율과 잘 관련지을 수 있어야 한다. 이러한 매개변수들을 잘 이해하면 분광 클래스는 적절한 정보 클래스로 명명될 수 있다.

원격탐사 자료를 판독하여 수집된 토지이용 및 토지 피복 자료를 쉽게 통합할 수 있도록 개발된 몇몇 범주형의 분류체계가 있다. 여기서는 다음의 몇 가지를 소

개한다.

- 미계획협회(American Planning Association : APA)의 상세한 토지이용 분류에 중점을 둔 토지기반 분류시스템(Land-Based Classification System : LBCS)(그림 9-3)
- 미지질조사국(United State Geological Survey : USGS)의 원격탐사 자료를 이용한 토지이용/토지피복 분류시스템(Land-Use/Land-Cover Classification System for Use with Remote Sensor Data)
- 미 어류 및 야생생물국(US department of the Interior Fish & Wildlife Service)의 습지와 심수 서식지 분류(Classification of Wetlands and Deepwaater Habitats of the United States)
- 미국가 식생 및 분류시스템(National Vegetation and Classification System)
- MODOS 토지피복 결과물을 만들기 위한 국제 지리 생물권 프로그램(International Geosphere-Biosphere Program : IGBP)의 토지피복 분류시스템(Land Cover Classification System) 수정본

#### 미계획협회의 토지기반 분류 표준

토지이용을 분류하기 위한 분류체계는 거의 없다. 사실상, 대부분의 사람들이 오직 토지피복 정보에만 관심이 있다고 말한다. 따라서 사용자가 비교적 고해상도 원격탐사 자료로부터 도심지/교외 토지이용 정보만을 추출하고자 한다면, 가장 실용적이고 포괄적인 계층적 분류 방법은 미계획협회(APA)(2004a)에서 개발한 토지기반 분류 표준(Land-Based Classification Standard : LBCS)이다. 이 표준은 표준 토지이용 코딩 안내서(Standard Land Use Coding Manual : SLUCM)를 갱신하는데(Urban Renewal Administration, 1965), SLUCM은 1987년 SIC (Standard Industrial Classification) 안내서(Bureau of the Budget, 1987)와 NAICS(North American Industrial

미계획협회의 토지기반 분류시스템(Land-Based Classification System)



표 1 : 사업

구획 ID	사업코드	설명
10-5-100	2100	쇼핑
10-5-100	2200	식당
10-5-100	6600	사교집회 및 종교집회
10-5-100	2100	가구
10-5-100	5210	주차장
기타	기타	기타

표 2 : 기능

기능코드	설명
2110	상업지구
2510	식당
6620	종교 기관
2121	가구
5200	주차장
기타	기타

표 3 : 건축물

구획 ID	건축물코드	설명
10-5-100	2500	몰, 쇼핑센터

그림 9-3 미계획협회는 도심지/교외 토지이용에 대해 세세한 정의를 담고 있는 토지기반 분류시스템(LBCS)을 개발하였다. 이 시스템은 현장 및 원격탐사 기법을 사용하여 추출된 정보를 활용한다. 이 그림은 CA주, Ontario의 경사항공사진이다. 이 지역에 대하여 가상적 사업 코드와 건축물 코드를 부여하였다. 지역개발과 소유주에 대한 정보 속성표는 나와 있지 않다 (American Planning Association 개념 참조).

Classification Standard) 개정판을 상호 참조한다.

LBCS는 아래와 같이 서로 다른 토지이용 분류시스템 간에 변환을 제공한다.

- SLUCM에서 LBCS로
- SIC에서 LBCS로
- NAICS에서 LBCS로

- 미 국방부 공군의 LCZC(Land-use Compatibility Zones Coding)에서 LBCS로

LBCS는 5가지 특성(사업, 기능, 지역개발, 건축물, 소유주)<sup>30)</sup>에 대한 정보를 필지 레벨에서 알기 위하여 현지 측량, 항공사진, 원격탐사 자료로부터 입력 자료를 구

30) activity, function, site development, structure, ownership

한다(American Planning Association, 2004b). 이 시스템은 모든 상업 및 공업지역의 토지이용에 대해 유일한 코드와 이에 대한 설명을 제공한다. LBCS 코딩을 주석으로 단 항공사진의 예가 American Planning Association (2004c)에 나와 있다. 그림 9-3은 CA주, Ontario 지역의 상가를 LBCS 사업, 기능, 건축물 코드로 표시한 예이다.

LBCS는 항상 개발 중에 있다. 정부에서는 사용자로 하여금 상업지와 공업지에 대한 토지이용 분류코드를 정확하게 알아야 하는 도심지/교외 연구에 LBCS를 이용하도록 독려하고 있다. LBCS는 도심지역에 있어서 토지피복이나 식생 특성을 제공하지 않는데, 이는 연방 지리자료위원회(Federal Geographic Data Committee : FGDC) 표준에 기초하고 있기 때문이다.

**USGS의 원격탐사 자료를 이용한 토지이용/토지피복 분류시스템**

USGS의 원격탐사 자료를 이용한 토지이용/토지피복 분류시스템(Anderson et al., 1976)은 자원중심의 토지피복 분류시스템이다. 이와 반대로 APA의 토지기반 분류시스템은 인간과 사업내역 중심의 토지이용 분류시스템이다. USGS의 이러한 자원중심의 분류 시스템은 “도시중심의 토지이용 분류시스템도 필요하지만 미국 토지의 남은 95% 자원을 대상으로 할 자원중심의 분류시스템도 필요하다”는 생각에 기초하고 있다. 이와 같은 필요성을 역설하기 위하여 USGS 시스템은 표 9-1에 표시한 바와 같이 레벨 I의 9개 항목 중에 도심지나 개발지역이 아닌 8개 항목을 다룬다. 이 시스템은 현장에서 얻어진 자료가 아닌 다양한 축척과 해상도로부터 얻어진 원격탐사 자료를 판독하여 추출하도록 설계되었다(표 9-2). 분류시스템은 계속해서 수정되고 있으며 USGS, 환경보호청(Environmental Protection Agency : EPA), NOAA 해안 서비스 센터(Coastal Service Center : CSC)와 다른 기관들에 의한 다양한 토지피복 매핑 활동을 지원하고 있다.

표 9-1 USGS의 원격탐사 자료를 이용한 토지이용/토지피복 분류시스템(Anderson et al., 1976).

분류레벨	
<b>1 도시 및 시가지</b>	
11	주거지
12	상업지
13	공업지
14	수송, 통신 및 공공시설
15	공업 및 상업 복합단지
16	혼합 도시 및 시가지
17	기타 도심지
<b>2 농업지</b>	
21	농경지 및 목초지
22	과수원 및 원예원
23	사육장(목장)
24	기타 농업지
<b>3 방목지</b>	
31	초본성 방목지
32	관목성 방목지
33	혼합 방목지
<b>4 산림지</b>	
41	낙엽수
42	상록수
43	혼합 산림
<b>5 수계</b>	
51	강 및 운하
52	호수
53	저수지
54	만 및 하구
<b>6 습지</b>	
61	산림습지
62	비산림습지
<b>7 나대지(불모지)</b>	
71	건염전
72	해안
73	해안 외 모래
74	노출된 암석
75	광산
76	전이지역
77	혼합 나대지
<b>8 툰드라</b>	
81	관목 툰드라
82	초본 툰드라
83	나지 툰드라
84	습지 툰드라
85	혼합 툰드라
<b>9 만년설 및 만년빙</b>	
91	만년설원
92	빙하

**표 9-2** USGS의 원격탐사 자료를 이용한 토지이용/토지피복 분류시스템에 대한 네 가지 레벨과 그 정보를 제공하기 위해 주로 사용되는 원격탐사 자료의 종류(Anderson et al., 1976 ; Jensen and Cowen, 1999).

분류 레벨	전형적인 자료 특성
I	NOAA AVHRR(1.1×1.1km), MODIS(250×250m ; 500×500m), Landsat MSS(79×79m), Landsat TM(30×30m), SPOT XS(20×20m) 와 같은 위성 영상
II	SPOT HRV 다중분광(10×10m) 및 인도 IRS 1-C 전정색(5×5m)과 같은 위성 영상. 1 : 80,000 축척보다 낮은 항공사진
III	1×1m~2.5×2.5m의 명목 공간해상도를 가진 위성영상. 중고도 항공사진(1 : 20,000 ~ 1 : 80,000 축척)
IV	1×1m 이하의 명목 공간해상도를 가진 위성 영상(예, QuickBird, IKONOS). 저고도 항공 사진(1 : 4,000 ~1 : 20,000 축척)

예를 들어, USGS의 토지피복 특성화 프로그램(Land-Cover Characterization Program : LCCP)은 토지피복 자료에 대한 국내외적인 요구에 의해 1995년에 시작되었으며 다음과 같은 일을 지원한다.

- 지구 토지피복 특성화 : 지구에 대한 1km 해상도의 토지피복 특성 데이터베이스를 개발하고자 하는 목적이 1997년에 달성되었다. 현재 세계 각국의 사용자들이 입력한 값을 사용하여 데이터베이스를 수정하는 데 주력하고 있다.
- 국가 토지피복 특성화 : 30m 해상도의 Landsat TM 영상 혹은 다른 자료를 이용한 NLCD(National Land-Cover Dataset)의 지속적인 제작

표 9-3은 USGS의 원격탐사 자료를 이용한 토지이용/토지피복 분류시스템이 어떻게 22개의 클래스가 되는지를 보여준다. 레벨 II 클래스의 대부분이 항공사진을 사용하여 잘 추출된다. 공간해상도와 영상판독에 대한 문제 때문에 Landsat TM 자료를 사용하여 레벨 II의 일부 클래스를 추출하는 것은 적합하지 않다. 따라서 Landsat

**표 9-3** 국가 토지피복 데이터셋과 NOAA 해안변화 분석 프로그램을 위해 수정된 USGS의 원격탐사 자료를 이용한 토지이용/토지피복 분류시스템(NOAA, 2004).

분류 레벨
<b>1 수계</b>
11 해수
12 만년빙/만년설
<b>2 개발지</b>
21 저밀도 주거지
22 고밀도 주거지
23 상업지/공업지/교통(운송)
<b>3 나대지</b>
31 노출된 암석/모래/점토
32 광산
33 전이지대
<b>4 육상산림</b>
41 낙엽수
42 상록수
43 혼합림
<b>5 관목지</b>
51 관목지
<b>6 인공삼림</b>
61 과수원 등
<b>7 초본성 육상식생</b>
71 초지/초본
<b>8 초본성 경작물</b>
81 목초/건초
82 열재배 농작물
83 작은 곡물
84 휴경지
85 도시/위락시설
86 초지
<b>9 습지</b>
91 산림습지
92 비영속성 초본습지

TM 자료를 사용하여 레벨 III 도심지 클래스와 같이 상세하게 분류하는 것은 시도되지 않았다. 또한 Anderson 레벨 II 클래스의 일부는 하나의 NLCD 클래스로 통합되었다(USGS, 2004).

해안변화 분석 프로그램(C-CAP) 결과물을 만들기 위해 NOAA 해안 서비스 센터에서도 동일한 클래스들을

사용한다. C-CAP 결과물에는 가능한 최근 토지피복 자료, 5년 전의 토지피복 자료 및 두 시기간의 변화를 설명하는 자료가 포함된다(Jensen et al., 1993 ; Dobson et al., 1995 ; NOAA, 2004).

연방 기관들이 모여 만든 MRLC(Multi-Resolution Land Characteristics) 협회는 미국의 Landsat 영상을 구입하고 토지피복 데이터셋을 만들기 위해 1992년에 결성되었다. 이 협회의 활동 중 하나가 2000년 9월에 1992년도 Landsat TM 자료로부터 Anderson et al.(1976)의 레벨 II 주제 정보와 동일한 범위를 가지는 NLCD를 완성한 것이다. 1999년에 2기 협회가 형성되어 전국적으로 일관성 있는 위성 원격탐사 및 토지피복 자료를 구축하고자 하는 연방 기관의 수요에 맞춰 MRLC 2001이라 불리는 국가 Landsat 영상 및 토지피복 데이터베이스를 구축하였다. MRLC 2001 영상 데이터베이스는 1780개에 해당하는 Landsat TM의 각 경로와 열 지역을 포함한다. MRLC 협회의 두 번째 목적은 미 50개 주와 Puerto Rico를 포함하는 NLCD 데이터베이스를 갱신하는 것이다. MRLC 2001은 1992년 NLCD와 양립될 수 있도록 표 9-3의 토지피복 분류시스템을 수정 갱신해 가고 있다.

#### 미 어류 및 야생 생물국의 습지와 침수 서식지 분류

미국은 농경지, 주거지, 상업지에 대한 지속적인 토지 이용 개발로 인해 내륙 및 해안 습지를 계속해서 잃어가고 있다. 미 어류 및 야생 생물국(US Department of the Interior Fish & Wildlife Service)은 미국의 습지를 지도로 만들고 이에 대한 목록을 만드는 곳이다. 따라서 원격탐사 자료와 현장 관측으로부터 추출된 정보를 활용하여 습지 분류시스템을 개발하였다(Cowardin et al., 1979). Cowardin 시스템은 생태학적 분류군을 정리하여 자원관리자가 통일된 용어와 개념을 가지고 시스템상에서 유용하게 그 정보를 이용도록 하였다. 습지는 식물의 특성, 지반, 범람의 정도에 따라서 분류된다.

생태학적으로 침수와 관련된 지역(예, 호수저층, 강바닥 등)은 일반적으로 습지로 고려되지 않고 침수 서식지로 분류된다. 해양, 하구, 강변, 호수, 늪지의 5개 시스템은 분류 계층에서 최상위 레벨에 있다(그림 9-4). 해양과 하구 시스템은 각각 2개의 하부 시스템(조하대와 조간대)을 가지고 있다. 강변 시스템은 4개의 하부 시스템(감조하천, 완경사 영구하천, 급경사 영구하천, 간헐적 하천)을 가지고 있다. 호수는 육수성 호수와 연안성 호수의 2개 하부 시스템을 가지고 있고 늪지는 하부 시스템이 없다. 하부 시스템 내의 클래스는 하층 물질의 종류, 범람 범위 또는 식생 형태에 따라 달라진다. 동일한 클래스가 하나 혹은 그 이상의 시스템이나 하부 시스템에 동시에 존재할 수 있다. 강변 시스템의 뚜렷한 특징은 그림 9-5에 나타낸 것과 같다. 이것은 국가적 차원에서 처음으로 인정된 습지 분류체계였다.

Cowardin 시스템은 FGDC의 습지위원회에 의해 습지 매핑과 목록을 구성하기 위한 국가 식물 분류 표준으로 채택되었다(FGDC, 1996). Cowardin 습지 분류시스템은 원격탐사 자료로부터 습지 정보를 추출하고 습지와 관련된 문제에 관심이 있는 사람들과 정보를 공유하려고 할 때 가장 실용적인 체계이다.

#### 미국가 식물 분류시스템

FGDC 산하 식물위원회는 국가 차원에서 통일성있는 식물 자원 자료를 얻기 위해 국가 식물 분류시스템(National Vegetation Classification System : NVCS)을 구축하였다(FGDC Vegetation Subcommittee, 1997 ; FGDC, 2004). NVCS는 자연적인 상태에서 군집으로 존재하는 식물을 분류하기 위해 체계적인 접근방법을 취하고 있다. 인상-식생학적(*physiognomic-floristic*) 분류체계는 다양한 축적에서 식생을 매핑하고 보존하는 데 적합한 정성적 및 정량적 자료를 사용한다. 표 9-4에서 볼 수 있듯이 인상학적 특성은 보다 일반적이고 덜 정밀한 수준의 분류를 제공하는 반면 식생학적 특성을 이용하면 보다 구

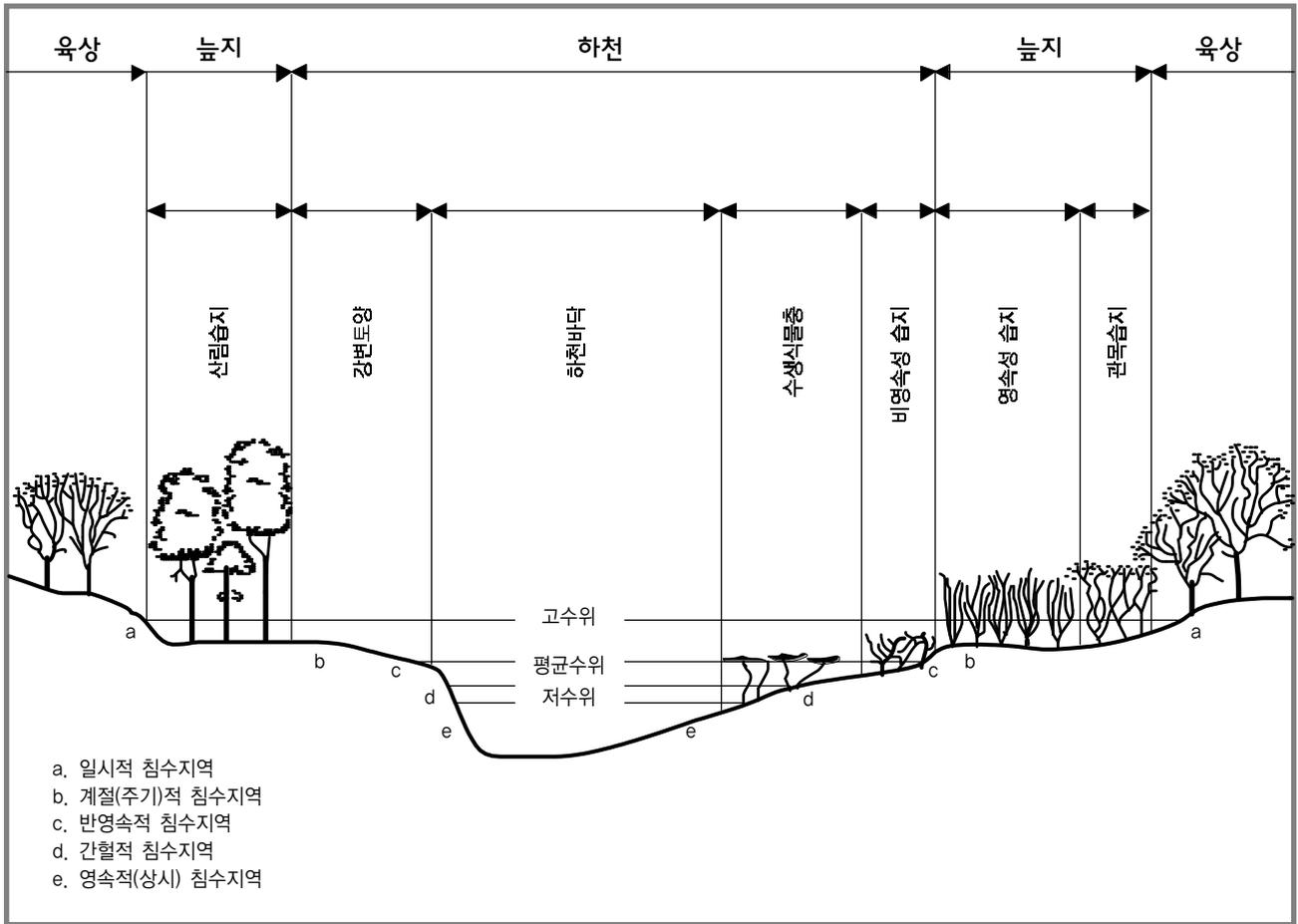


그림 9-4 미국의 습지와 침수 서식지 분류에서 강변 시스템 상의 뚜렷한 특징과 서식지(Cowardin et al., 1979).

체적인 수준의 분류체계가 가능하다.

현재, 인상학적 분류를 위한 기준은 잘 정립되어 NVCS에 포함되어 있으나(ESA, 2004a) 식생학적 분류 기준은 일부분만 개발되어 있다. 미국 생태학회(Ecological Society of America : ESA)의 식물 분류 위원회는 “미국 식생의 표준화된 분류”라는 식생학적 표준을 개발하고 있다(ESA, 2004b). 만약 과학자가 원격탐사 자료로부터 상세한 식생 정보를 추출하고 이를 생태학적으로 안정된 식생 분류시스템에 위치하는데 관심이 있다면 국가 식물 분류시스템이 적합할 것이다.

### MODIS 토지피복 결과물을 만들기 위한 IGBP 토지 피복 분류시스템 수정본

지역적, 국가적, 그리고 전지구적 규모의 토지피복 정보를 구축하고자 한다면 IGBP 토지피복 분류시스템 수정본이 적합할 것이다(Loveland et al., 1999). 예를 들어, NASA EOS의 MODIS는 지구의 토지표면 정보를 250m에서 1,000m의 공간해상도로 제공한다. 지구적 변화를 연구하기 위해 사용하는 표준 MODIS 결과물이 대략 44개가 있다. 예를 들어, MODIS 토지 과학 팀은 큰 규모의 토지피복 변화를 설명하기 위해 1km(0.6마일) 해상도로 전지구적 규모의 토지피복 변화지도를 만든다.

토지피복 종류와 토지피복 변화를 나타내는 변수들

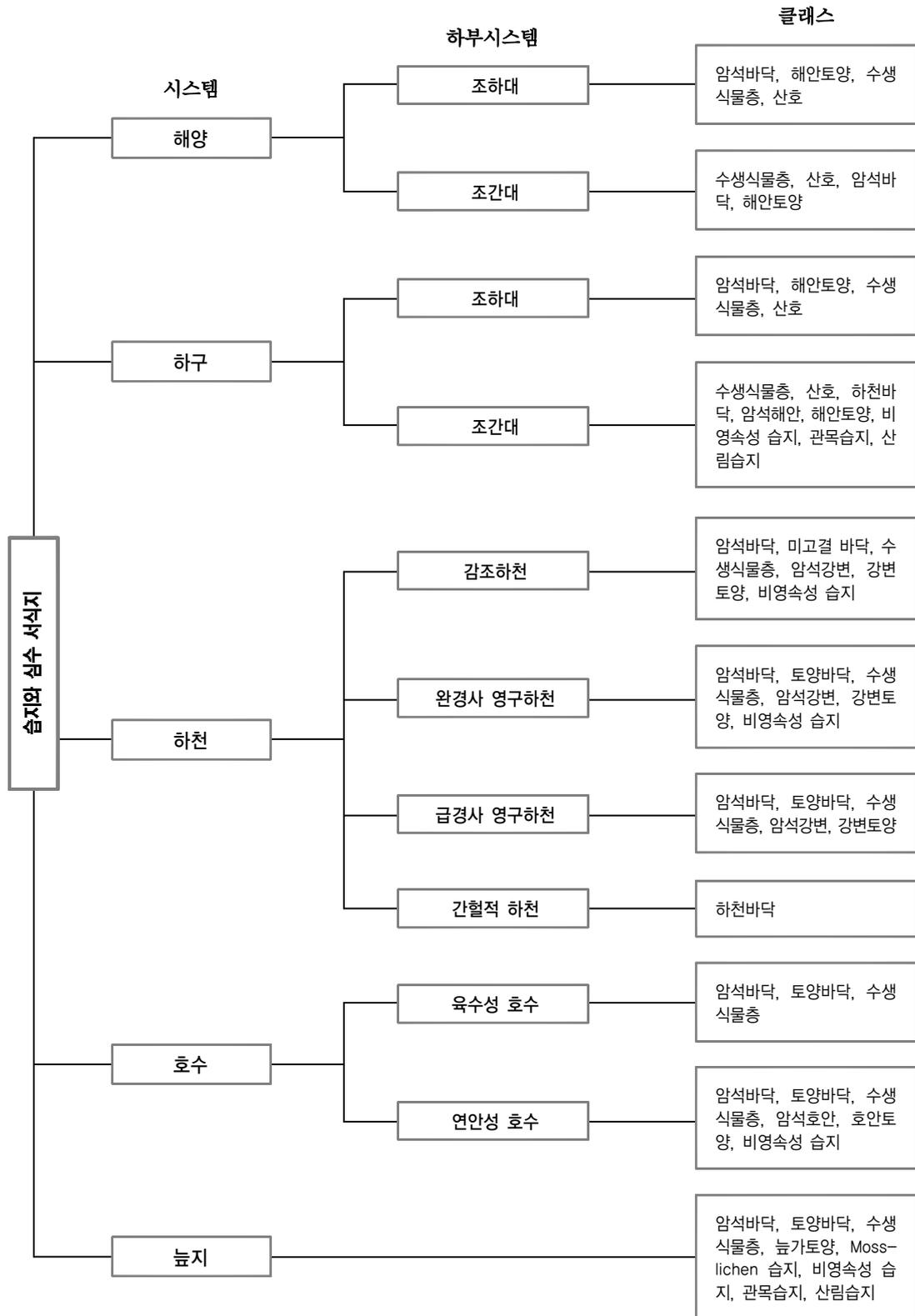
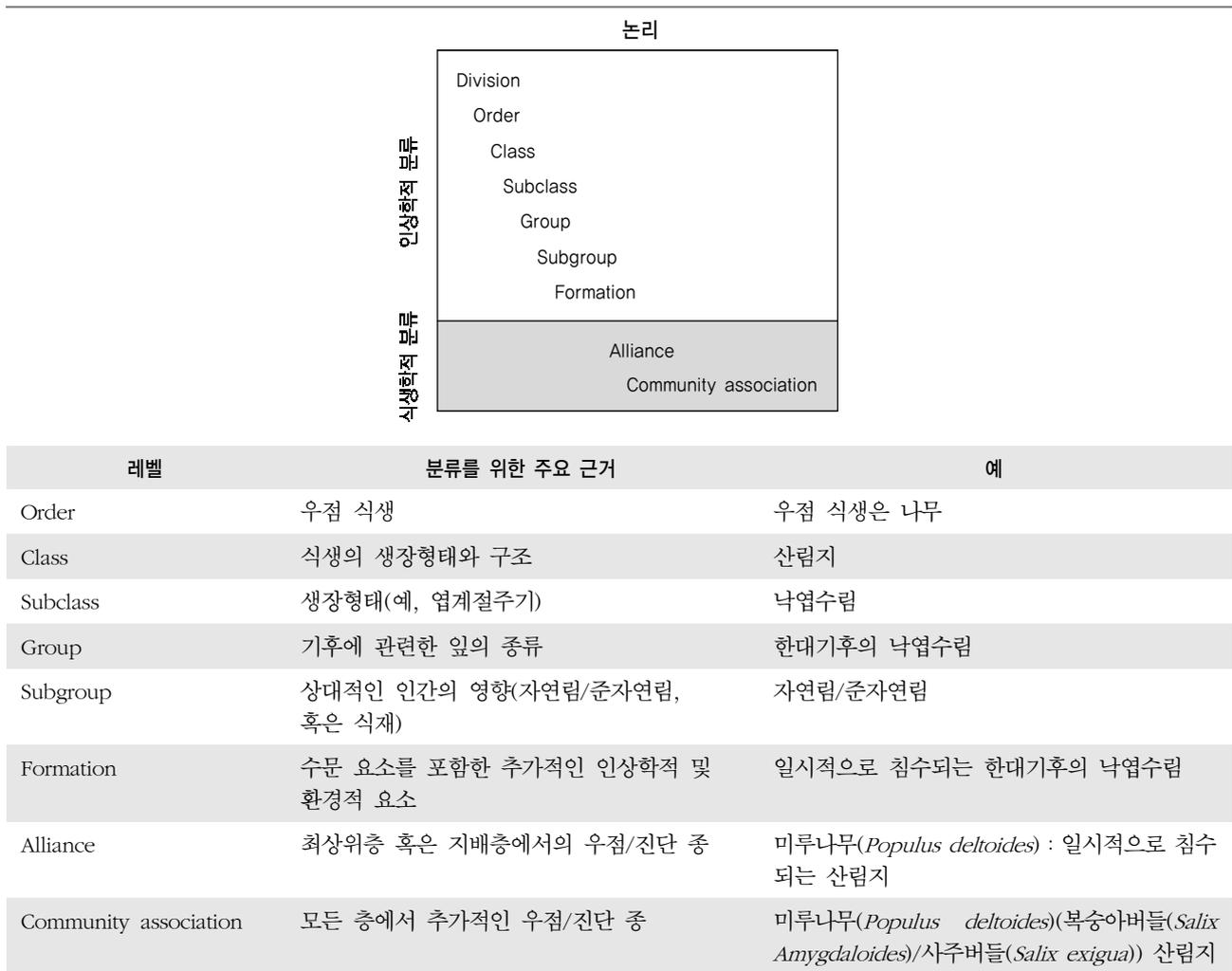


그림 9-5 미국의 습지와 심수 서식지 분류의 습지 서식지 시스템, 하부시스템, 그리고 클래스의 분류체계(Cowardin et al., 1979). 늪지 시스템은 심수 서식지를 포함하지 않는다. Cowardin 시스템은 습지 매핑과 목록을 위한 국가 식물 분류 기준이다 (Wetlands Subcommittee of the Federal Geographic Data Committee ; FGDC, 2004).

표 9-4 국가 식물 분류시스템의 개념, 레벨 및 주요 기준(FGDC Vegetation Committee, 1997 ; FGDC, 2004).



은 1km 해상도로 1년에 4번 만들어진다. 토지피복 변수를 이용해 IGBP 식물 데이터베이스를 따르는 17가지의 토지피복 항목을 식별할 수 있다(표 9-5). IGBP 식물 데이터베이스는 자연 식물에 대하여 9개 클래스, 개발지에 대하여 3개의 클래스, 혼합 토지 지역에 대하여 2개 클래스, 비식생 토지(눈/얼음, 나대지/암석, 물) 지역에 대하여 3개 클래스로 나누어 정의한다(Strahler, 1999 ; NASA GSFC, 2004). MODIS 자료에 기초하여 만든 첫 번째 토지피복도는 2002년 8월에 배포되었다. 1km 해상도의 이 지도는 2000년 11월과 2001년 10월 사이에 수집된 MODIS 영상을 기반으로 만들었다

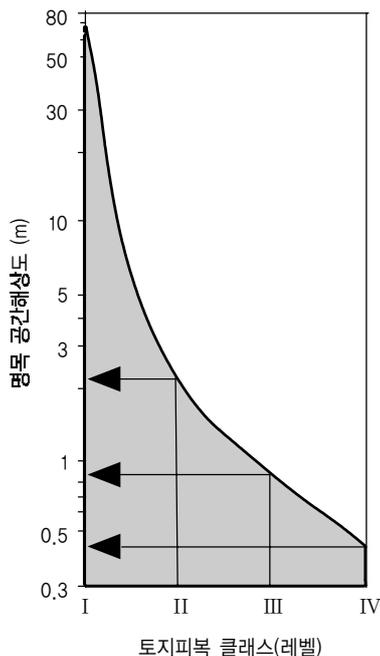
(NASA Earth Observatory, 2002). MODIS를 이용해 만든 북미지역의 토지피복도가 **컬러도판 9-1**에 나와 있다(Friedl et al., 2002).

1km 해상도의 기본적인 결과물 이외에도, 토지피복의 비율과 변화 특성을 포함한 요약 결과물도 1/4도, 1/2도, 1도 해상도로 이용가능하다. 이러한 결과물들은 250m 해상도의 토지피복 변화지도와는 독립적으로 준비되고 사용되는 알고리즘도 다르다(EROS, 2003). 인간활동에 의해 야기된 토지피복 변화는 일반적으로 1km 해상도보다 더 높은 해상도에서 잘 감지되는데, Maryland 대학은 이러한 변화를 보기 위해 250m 해상

표 9-5 MODIS 토지피복 결과물을 만들기 위한 IGBP 토지피복 분류시스템의 수정본(Strahler, 1999 ; Loveland et al., 1999 ; Friedl et al., 2002 ; NASA Earth Observatory, 2002).

자연식생	
<b>산림</b>	
상록 침엽수림	수목식생이 60% 이상을 차지하며 높이가 2m 이상의 나무가 주를 이룸. 대부분의 나무는 연중 녹색을 띠며, 임관은 녹엽을 항상 가지고 있다.
상록 활엽수림	수목식생이 60% 이상을 차지하며 높이가 2m 이상의 나무가 주를 이룸. 대부분의 나무와 관목은 연중 녹색을 띠며, 임관은 녹엽을 항상 가지고 있다.
낙엽 침엽수림	수목식생이 60% 이상을 차지하며 높이가 2m 이상의 나무가 주를 이룸. 계절적 침엽수 군집은 매년 잎이 있는 시기와 없는 시기가 반복된다.
낙엽 활엽수림	수목식생이 60% 이상을 차지하며 높이가 2m 이상의 나무가 주를 이룸. 활엽수 군집은 매년 잎이 있는 시기와 없는 시기가 반복된다.
혼합림	나무식생이 60% 이상을 차지하며 높이가 2m 이상의 나무가 주를 이룸. 여러 산림 유형이 섞여 있음. 어떠한 산림 유형도 60% 이상을 차지하지 못한다.
<b>관목, 초지 및 습지</b>	
폐쇄형 관목지	2m 이하의 수목으로 구성되어 있으며 관목 임관이 60% 이상. 관목은 상록수이거나 낙엽수이다.
개방형 관목지	2m 이하의 수목으로 구성되어 있으며 관목 임관이 10~60%임. 관목은 상록수이거나 낙엽수이다.
수목이 우거진 사바나	초본과 기타 하층 시스템으로 30~60%의 산림 임관이 존재. 산림은 2m 이상의 수목으로 구성
사바나	초본과 기타 하층 시스템으로 10% 미만의 산림 임관이 존재. 산림은 2m 이상의 수목으로 구성
초지	초본류 지역으로 나무와 관목은 10% 미만임
영구습지	물과 초본류 혹은 나무 식생의 혼합으로 영구적인 지역. 식생은 염수, 기수, 담수에 존재할 수 있음
<b>개발지 및 혼합지</b>	
<b>농업지</b>	
농작지	일시적인 농작물에 의해 덮여 있다가 추수 후 나지로 변하는 주기적인 지역(예, 단일 및 다중 농작 시스템). 연중 나무 농작물은 적절한 산림이나 관목 유형으로 분류된다.
<b>농작물/자연식생 혼합지</b>	
농작물/자연식생 혼합지	농작지, 산림, 관목지, 초지 등의 혼합지로 어떠한 유형도 전체의 60% 이상을 넘지 못한다.
<b>도시</b>	
시가지	건물과 기타 인간에 의해 만들어진 구조물 지역
<b>비식생 지역</b>	
<b>나대지</b>	
나대지 혹은 식생이 드문 지역	토양, 모래, 암석, 혹은 눈. 연중 식생에 의해 덮여 있는 면적이 10%미만인 지역
<b>설빙</b>	
설빙	연중 내내 설빙으로 덮여 있는 지역
<b>수계</b>	
수계	해양, 호수, 저수지, 강. 담수 혹은 염수

도에서 결과물을 제공하고 있다(Townshend, 1999). Zhang et al., 2003).  
 구의 식물 계절학 정보 또한 이용 가능하다(Zhang et



**그림 9-6** 미국 토지피복 클래스의 레벨 I부터 레벨 IV를 매핑하기 위해 필요한 최소 공간해상도(Anderson et al.,(1976)을 기반으로 함). 레벨 II 클래스를 제작하기 위해서 요구되는 공간해상도가 급격히 증가한다는 사실에 주목할 것(Welch, 1982 ; Jensen and Cowen, 1999).

**분류체계에 대한 고찰**

원격탐사 자료와 같은 지리 정보는 종종 부정확하다. 예를 들어, 산림과 방목지 사이는 점진적으로 변하는데 비하여 앞에서 언급된 분류 개념은 이러한 전이지대의 클래스들 간에 범주형의 경계가 필요하다. 그러나 여기에 적용되는 분류체계는 주제 정보가 퍼지하므로 퍼지 정의를 포함해야 한다(Wang, 1990a). 퍼지 분류 개념은 일반적으로 표준화되어 있지 않으며 개개 연구자들에 의해 특정 지역의 프로젝트에서 주로 개발되었다. 이렇게 개발된 퍼지 시스템은 다른 환경에서는 적용 못할 수도 있다. 따라서 우리는 현존하는 범주형 분류체계의 사용을 주로 살펴볼 것인데, 이러한 분류체계는 고정적이고 선형적 지식에 기초하며, 일반적으로 사용하기 어렵다. 이러한 분류체계는 과학적이기 때문에 계속 널리 사용할 수 있고 서로 다른 개인이 동일한 분

류체계를 사용하여 결과를 비교할 수 있다.

만약 이미 뛰어난 분류시스템이 존재하면, 제한적으로만 사용될 수 있는 완전히 새로운 시스템을 개발하는 것은 어리석은 일이다. 국내외적으로 이미 사용하고 있는 분류시스템을 채택하거나 수정하는 것이 더 낫다. 이렇게 함으로써 다른 연구의 관점에서 분류결과를 비교 해석할 수 있고 자료를 공유하는 것이 더 쉬워질 것이다.

마지막으로, 분류체계에서의 상세도 수준과 정보를 제공하기 위하여 사용되는 원격탐사 영상의 해상도 사이에는 일반적인 관계가 있다는 사실을 기억해둘 필요가 있다. Welch(1982)는 미국의 도심/교외 토지이용도와 토지피복도를 만드는데 이러한 관계를 요약하였다(그림 9-6). 식생 지도를 제작할 때도 비슷한 관계가 성립한다(Botkin et al., 1984). 예를 들어, 센서 시스템과 식생을 구별하는 데 유용한 공간해상도 사이의 관계를 광역의 관점에서부터 국지적인 관심에서 본 결과를 그림 9-7에 요약하였다. 이 그림은 분류시스템상의 상세도 수준에 따라 사용되어야 하는 원격탐사 자료의 공간해상도를 나타내고 있다. 물론, 원격탐사 시스템의 분광 해상도 역시 고려해야 하는데, 특히 식물, 물, 얼음, 눈, 토양, 암석 등을 조사할 때에는 반드시 고려해야 한다.

**훈련지역 선택 및 통계값 추출**

영상 분석기는 분류체계를 선택한 후 대상 토지피복 클래스나 토지이용 클래스를 대표할 수 있는 영상 안의 훈련지역을 선택할 것이다. 훈련자료는 수집된 환경이 비교적 균일할 때 그 가치가 있다. 예를 들어, 초원지대에 속한 모든 토양이 배수가 잘되는 비옥한 토양이라면 그 지역에서 선택된 초지 훈련자료는 좋은 자료가 될 수 있다. 그러나 대상지역의 토양 속성이 급격히 변한다면(예, 대상지역의 절반이 표층 근처에 주수 지하수면을 가지고 있어 수분이 많은 경우) 대상지역의 건조

**레벨 I : 전지구**

AVHRR  
MODIS  
해상도 : 250m~1.1km



**레벨 II : 대륙**

AVHRR  
MODIS  
Landsat MSS  
Landsat TM  
해상도 : 80m~1.1km



일반화된  
식생분류

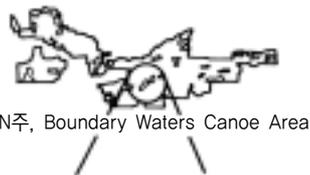
**레벨 III : 생물군계**

Landsat MSS  
Landsat TM  
ETM<sup>+</sup> SAR  
해상도 : 30m~80m



**레벨 IV : 지역**

Landsat TM  
SPOT  
높은 고도에서 촬영한 항공사진  
SAR  
해상도 : 3m~30m



**레벨 V : 도화**

입체항공사진  
IKONOS  
QuickBird  
해상도 : 0.25m~3m



**레벨 VI : 현장 측량**

표면 측량 및 관측



그림 9-7 식생 목록을 위한 대표적 원격탐사 시스템의 공간해상도와 필요한 상세 레벨 사이의 관계.

한 토양부분에서 선택된 초지 훈련자료는 수분함량이 많은 토양부분의 초지의 분광상태를 대표하지 못할 것이다. 따라서 지리적 확장(*geographic signature extension*) 문제가 발생하는데, 이는 원격탐사 훈련자료를  $x, y$  공간으로 확장하는 것이 가능하지 않을 수도 있다는 것을 뜻한다.

이러한 상황을 해결하는 가장 쉬운 방법은 프로젝트의 준비 단계에서 지리적 층화(*geographical stratification*)를 적용하는 것이다. 이때 지리적 확장 문제를 야기하는 중요한 환경적 요인들을 정확하게 파악하여야 한다. 환경적 요인이란 토양 종류, 물 혼탁도, 농작물 종류, 심한 폭우에 의해 생긴 토양의 이상 수분함량, 대기 중에 국부적으로 존재하는 연무 등의 차이와 같은 것이다. 영상에서 이러한 환경적 요인들이 작용하는 것을 분명히 밝혀서 지역적으로 구분할 수 있는 훈련지역을 선택하는 것이 바람직하다. 이러한 경우, 비교적 짧은 거리상에서 분류자를 훈련시켜야 할 수도 있으며, 개개의 층에 대하여 독립적으로 분류를 실시해야 할 수도 있다. 따라서 전체 지역에 대한 분류지도는 개개 층에 대하여 얻은 결과물을 조합하는 형태가 된다. 그러나 만약에 밴드비율 영상이나 대기보정 등을 통해 환경적 요소가 균일하거나 일정하게 유지될 수 있다면, 훈련시 필요한 비용과 노력을 대폭 감소시키며 훈련지역을 공간상에서 원거리까지 확장하여 적용할 수 있을 것이다. 훈련지역을 시공간적으로 확장하는 개념을 완전히 이해하기 위해서는 다른 부수적인 연구가 필요하다.

일단 시공간 확장요인이 고려되면 분석가는 각 클래스를 대표할 수 있는 훈련지역을 선택하고 각 훈련지역 내의 화소의 분광 통계값을 수집한다. 일반적으로 훈련자료는 많은 화소로 구성되어 있다.  $n$ 밴드로부터 훈련자료가 추출되었다면 각 클래스에 대하여 훈련자료의  $10n$  이상의 화소를 수집하는 것이 일반적인 규칙이다. 이는 일부 분류 알고리즘에서 필요로 하는 분산-공분산 행렬을 계산하는 데 충분하다.

훈련지역 자료를 수집하는 데 다음과 같은 여러 가지 방법이 있다.

- 나무 종류, 나무 높이, 임관 백분율, 흉고직경(*diameter-at-breast-height : dbh*) 측정과 같은 현장 정보의 수집
- 화면상에서의 폴리곤 형태의 훈련자료 선택
- 화면상에서 초기 훈련자료 할당

이상적으로, 각각의 훈련지역은 직접 방문하여 훈련지역의 주변길이 혹은 중심좌표는 GPS 기기를 사용하여 직접 관측한다. 미국 정부가 선택적 가용성(*Selective Availability : SA*)을 제거하였을 때 보정된 GPS(즉, DGPS)의 수평 ( $x, y$ ) 좌표는 오차범위  $\pm 1m$  안에 있어야 한다. GPS로 수집한 훈련지역(예, 오크나무 숲을 둘러싸고 있는 폴리곤)의 ( $x, y$ ) 좌표는 영상처리 시스템으로 바로 입력될 수 있고 훈련 클래스(오크 산림) 통계를 추출하는 데 사용될 수 있다.

분석가는 영상을 컬러 CRT 화면을 통해 보면서 폴리곤 형태의 관심영역(*Area Of Interest : AOI*)을 선택할 수 있다. 대부분의 영상처리 시스템은 상세한 AOI를 인식할 수 있도록 폴리곤을 마음대로 그릴 수 있는 고무 밴드(*rubber band*) 툴을 사용한다. 반대로, 커서를 사용해서 영상 공간의 정확한 위치를 선택할 수도 있다. 시드(*seed*) 프로그램은 1개의 ( $x, y$ ) 위치에서 시작하여 이웃한 화소의 값을 모든 밴드에 대하여 평가한다. 분석가가 정한 기준을 사용하여 시드 알고리즘은 원래의 시작점의 화소와 비슷한 분광특성을 가진 화소를 찾지만 하면 아메바처럼 확장해 나간다. 이 알고리즘은 균일한 훈련 정보를 찾는 데 매우 효과적인 방법이다.

특정한 클래스 ( $c$ )와 연관된 각각의 훈련 지역의 화소는 다음의 측정벡터  $X_c$ 로 나타낼 수 있다.

$$X_c = \begin{bmatrix} BV_{i1} \\ BV_{i2} \\ BV_{i3} \\ \vdots \\ BV_{ik} \end{bmatrix} \tag{9-1}$$

여기서  $BV_{i,j,k}$ 는 밴드  $k$ 의  $(i, j)$ 번째 화소의 밝기 값이다. 각 밴드의 훈련 클래스에 해당하는 화소의 밝기 값을 통계적으로 분석하여 각 클래스에 대한 다음의 평균값 벡터  $M_c$ 를 산출한다.

$$M_c = \begin{bmatrix} \mu_{c1} \\ \mu_{c2} \\ \mu_{c3} \\ \vdots \\ \mu_{cn} \end{bmatrix} \quad (9-2)$$

여기서  $\mu_{ck}$ 는 밴드  $k$ 의 클래스  $c$ 에 대해 수집된 자료의 평균값을 뜻한다. 원시 측정벡터를 분석하여 각 클래스  $c$ 에 대한 공분산 행렬을 만들 수 있다.

$$V_c = V_{ckl} = \begin{bmatrix} Cov_{c11} & Cov_{c12} & \cdots & Cov_{c1n} \\ Cov_{c21} & Cov_{c22} & \cdots & Cov_{c2n} \\ \vdots & \vdots & \ddots & \vdots \\ Cov_{cn1} & Cov_{cn2} & \cdots & Cov_{cnn} \end{bmatrix} \quad (9-3)$$

여기서,  $Cov_{ckl}$ 는 밴드  $k$ 와 밴드  $l$  사이의 클래스  $c$ 에 대한 공분산 값이다. 클래스  $c$ 의 공분산 행렬( $V_{ckl}$ )의 기호는 간단히  $V_c$ 로 줄여서 쓰기로 한다. 클래스  $d$ 의 공분산 행렬도 마찬가지다( $V_{dkl} = V_d$ ).

SC주, Charleston 지역에 대한 5가지의 토지피복 클래스(주거지, 상업지, 습지, 산림, 수계)의 훈련지역에 대한 평균, 표준편차, 분산, 최소값, 최대값, 분산-공분산 행렬, 상관 행렬이 표 9-6에 나와 있다. 이것은 5개 클래스의 기본적인 분광 특성정보를 나타낸다.

수동적으로 훈련지역 폴리곤을 선택하면 훈련 클래스 히스토그램상에서 다중 모드를 가지기도 하는데, 이는 해당 훈련지역 내에 적어도 두 종류 이상의 토지피복이 존재한다는 것을 뜻한다. 이러한 상황은 개개 클래스를 구별하려고 할 때 좋지 않기 때문에 다중 모드 훈련자료는 버리거나 단일모드 히스토그램을 가지도록 폴리곤을 줄여서 사용하도록 한다.

연속되거나 이웃하는 화소들 간에는 양의 자기상관성(*autocorrelation*)이 존재한다(Gong and Howarth, 1992).

이것은 인접한 화소는 비슷한 밝기값을 가질 확률이 높다는 것을 의미하는데, 자기상관성이 높은 자료에서 수집된 훈련자료는 분산이 감소하는 경향이 있다. 이는 실제 현장 상태에 의해서라기보다 센서가 자료를 수집하는 방법에 따라 더 영향을 받는다(예를 들어, 대부분의 감지기는 각각의 화소에 매우 짧은 시간 동안 머물기 때문에 해당 화소와 이웃한 화소의 분광정보를 완전히 분리시키지 못할 수도 있다). 이상적인 조건은 매  $n$ 번째 화소를 사용하거나 다른 표집 기준을 사용하여 연구지역의 훈련자료를 수집하는 것이다. 이는 자기상관되지 않은 훈련자료를 수집하기 위함이나, 대부분의 디지털 영상처리 시스템은 훈련자료 수집 모듈에서 이러한 옵션을 제공하지 않는다.

### 영상 분류를 위한 최적 밴드 선택 : 피쳐선택

관심 클래스에 대해 각 밴드의 훈련 통계값이 체계적으로 수집되면 한 클래스를 다른 클래스와 구별하기에 가장 효율적인 밴드들을 결정해야 하는데, 이 과정을 보통 피쳐선택(*feature selection*)이라고 부른다(Duda et al., 2001). 분석할 때 불필요한 분광 정보를 제공하는 밴드를 제거하는 것이 목적이며, 이러한 방법으로 데이터셋의 차원(분석할 밴드 수)을 줄일 수 있다. 이러한 과정은 디지털 영상 분류 과정에 소요되는 비용을 최소화하나, 이로 인해 정확도가 떨어지는 안 된다. 피쳐선택을 하려면 원격탐사 훈련자료 내에서 클래스간 분리를 결정해야 하는데, 여기에는 통계적 분석과 그래프 분석 방법이 포함된다. 통계적인 방법을 이용할 경우, 밴드조합은 일반적으로  $n$ 밴드를 동시에 사용하여 각 클래스를 다른 클래스로부터 구별할 수 있는 잠재적인 능력에 따라 순위가 매겨진다(Beauchemin and Fung, 2001). 발산(*divergence*)과 같은 통계 측정값에 대하여 곧 다룰 것이다.

만약 통계적 접근 방법이 분류에 가장 적절한 밴드를