공 학 석 사 학 위 논 문

고해상도 위성영상을 이용한 도심지 토지이용 분석

2006년 2월

부 경 대 학 교 대 학 원

토 목 공 학 과

김 세 준

공 학 석 사 학 위 논 문

고해상도 위성영상을 이용한 도심지 토지이용 분석

지도교수 이 종 출

이 논문을 공학석사 학위논문으로 제출함

2006년 2월

부 경 대 학 교 대 학 원

토 목 공 학 과

김 세 준

김세준의 공학석사 학위논문을 인준함

2006년 2월 24일

주 심	공학박사	이 환 우	(인)
위 원	공학박사	정 두 회	(인)
위 원	공학박사	이 종 출	(인)

목 차

List of Tables i
List of Figures iii
Notation iii

Abstract iv

1. 서 론 1

1.1 연구 배경 및 목적 1
1.2 연구동향 2
1.3 연구방법 및 범위 4

2. 연구의 이론적 배경 7

2.1 위성영상 전처리 7
2.1.1 방사보정 7
2.1.2 기하보정 7
2.2 영상 분류 이론 10
2.2.1 감독분류 11
2.2.1.1 평행육면체분류 12
2.2.1.2 최대우도분류 13
2.2.1.3 최소거리분류 14
2.2.2 무감독분류 15

2.2.3 신경망분류 ····· 17
2.3 분류영상의 정확도 검증기법 ····· 19
2.3 GIS 데이터 모델 ····· 23
2.3.1 공간데이터베이스 ····· 24

3. 분광특성을 이용한 위성영상 분류 ····· 26

3.1 연구지역 ····· 26
3.2 위성영상의 전처리 ····· 27
3.3 Test 지역 선정 ····· 29
3.4 감독분류 ····· 30
3.5 무감독분류 ····· 32
3.6 신경망분류 ····· 37
3.7 분류 결과 분석 ····· 39

4. GIS를 이용한 토지피복도 작성 ····· 41

3.1 토지피복도 작성 방법 ····· 43
3.2 GIS를 이용한 토지피복도 작성 ····· 44
5. 결 론 ····· 47

참고문헌 ····· 48

List of Tables

<Table 2.1> Specific of Classification ····· 11

<Table 2.2> Error Matrix ····· 19

<Table 3.1> Specific of IKONOS OSA ····· 26

<Table 3.2> Control Point of RMSE ····· 28

<Table 3.3> Covariance Matrix(Supervised Classification) ····· 31

<Table 3.4> Univariate Matrix(Supervised Classification) ····· 32

<Table 3.5> Error Matrix(Supervised Classification) ····· 32

<Table 3.6> Covariance Matrix(K-means) ····· 35

<Table 3.7> Univariate Matrix(K-means) ····· 35

<Table 3.8> Error Matrix(K-means) ····· 35

<Table 3.9> Covariance Matrix(ISODATA) ····· 36

<Table 3.10> Univariate Matrix(ISODATA) ····· 36

<Table 3.11> Error Matrix(ISODATA) ····· 36

<Table 3.12> Covariance Matrix(Neural Network) ····· 38

<Table 3.13> Univariate Matrix(Neural Network) ····· 39

<Table 3.14> Error Matrix(Neural Network) ····· 39

<Table 3.15> Classification of Overall Accuracy, Kappa ····· 41

<Table 3.16> Kappa Quality ····· 41

<Table 3.17> Number of Class Pixel ····· 41

List of Figures

<Fig. 1.1> Flow Chart of Research Planning ····· 6

<Fig. 2.1> Geometric Transformation ····· 8

<Fig. 2.2> Parallel-Piped Classifier ····· 13

<Fig. 2.3> Maximum Likelihood ····· 14

<Fig. 2.4> Minimum Distance Classification ····· 15

<Fig. 2.5> GIS Data Model Notion ····· 23

<Fig. 2.6> Conversation of Data from Raster to Vector Format 25

<Fig. 3.1> IKONOS Image of Study Area ····· 27

<Fig. 3.2> Geometric Correction ····· 28

<Fig. 3.3> Test Area ····· 29

<Fig. 3.4> Land Cover Test Area Using Supervised Classification 31

<Fig. 3.5> Land Cover Test Area Using K-means ····· 34

<Fig. 3.6> Land Cover Test Area Using ISODATA ····· 34

<Fig. 3.7> Land Cover Test Area Using Neural Network ····· 38

<Fig. 4.1> Flow Chart of Land-Cover Map By GIS ····· 43

<Fig. 4.2> Allowance of Spatial and Attribute Data on GIS ····· 44

<Fig. 4.3> DEM GIS Database ····· 45

<Fig. 4.4> Land Cover Classification Map(Supervised Classification) 45

<Fig. 4.5> Land Cover Classification Map(Neural Network) ····· 46

Notations

X_i	:	입력자료
N_l	:	입력 자료의 차수
N_H	:	은닉층 Node의 수
θ_j	:	은닉층 j번째 Node사이의 가중치
w_{ij}	:	Node 사이의 가중치
N_0	:	출력층 Node의 수
Q_k	:	출력층 k번째 주 Node의 Bias

An Analysis of Land Use in Urban Area Using High-resolution Satellite Image

Se - Jun, Kim

Department of Civil Engineering, Graduate School,
Pukyong National University

ABSTRACT

We need the acquisition of accurate geographic information as well as immediate updates of information on the city in order to plan and manage the changes of the cities more systematically. The geographic information for judging the changes of the cities can be used not only in various policies and studies, but also as important data themselves that record the growth of cities. The variety of methods for obtaining of this geographic information and wide-spread changes of urban areas are being presented, the use of the aerial photograph or the satellite image of them enables us to obtain and update The geographic information immediately. Since especially the *H*igh-Resolution satellite image includes existing character of aerial photograph as well, the study on this subject is progressing vigorously.

The IKONOS satellite image which was launched recently in 1999 provides the spatial resolution that can solve the problem caused by the impropriety of the existing resolution, especially We can say that the multi spectral image is the appropriate image for the analysis of land cover.

In this paper, therefore, the land cover classification that took advantage of spectral character with the IKONOS image suggested that the accuracy of the supervised classification that made use of the Maximum likelihood method was the highest accuracy, and the classification method that made use of the

neural network had the highest accuracy in the classification method with subjectivity was not to be considered. By using the land cover map classified by each classification method, we could build the GIS database with attribute data about the classification accuracy and the class ,and by providing the land cover map by each classification method according to the accuracy that the user requires, we could provide more reliable and various information.

1. 서 론

1.1 연구 배경 및 목적

오늘날 경제성장과 인구 집중 현상으로 도시지역의 발전은 기존 도시를 외곽으로 확대시켰으며, 도시 내 환경을 급속히 변화 시키고 있다. 이러한 도시의 변화에 대하여 보다 체계적으로 계획하고 관리하기 위해서는 도시지역에 대한 정확한 지리 정보의 획득이 필요하며 이와 더불어 정보의 신속한 갱신이 필요하다. 이러한 도시 변화를 판단하기 위한 지리정보는 여러 가지 정책과 연구에 사용될 수 있을 뿐만 아니라 그 자체만으로도 도시의 성장을 기록하는 중요한 자료로 이용될 수 있다. 이러한 지리정보의 획득과 광범위한 도시지역의 변화를 추출하기 위해서는 다양한 방법이 제시되고 있으며, 그 중 하나는 지형도를 비교하여 도시지역의 변화를 추출하는 것이고, 다른 하나는 항공사진이나 위성영상 등과 같은 영상데이터를 이용하는 방법이다[1]. 항공사진이나 위성영상의 사용은 공간 정보를 짧은 시간에 획득할 수 있고 빠른 시기에 갱신할 수 있다. 특히 고해상도 위성영상은 기존의 항공사진이 갖고 있던 특성까지도 포함하고 있어 연구가 활발하게 진행되고 있다.

그러나 고해상도 위성영상은 항공사진과 달리 기하학적 특성이 다르고 모델링에 따르는 다양한 방법을 적용할 수 있는 반면 정확도를 만족하기 위한 다양한 사전 연구를 수행하여야 한다. 그 중 위치정확도와 관련해서 정량적 해석이외에 실제 지형물의 판독 가능성에 대한 분석이 요구되지만 해상력의 부적절로 인하여 개력적인 토지이용현황을 파악하는데 그쳤다. 그러나 도시지역은 외곽 지역과 달리 주택, 교통시스템, 도시시설, 상업용 건물 그리고 여가공간 등 복잡한 형태의 물질들로 구성되어 있기 때문에 토지의 분류에 있어 위성영상은 높은 공간 해상도가 요구된다[2].

1999년도에 발사한 IKONOS 위성영상은 흑백 및 다중분광 영상 모두 11비트의 정보를 가지고 있으며, 1m급 전정색(panchromatic) 영상은 가시광선에서 근적외선 파장대의 정보를 가지고 있다. 4m급 다중분광(multispectral) 영상은 적색, 녹색, 청색과 근적외선의 각 4개의 밴드를 제공한다. 이와 같이 IKONOS영상은 도시지역의 토지피복 환경 분석에 있어 필요한 공간 해상력을 제공하고 있으며, 다중분광 영상은 토지 피복의 분석에 있어 적합한 영상이라고 할 수 있다[3].

따라서 본 연구에서는 고해상도 IKONOS 위성영상을 이용하여 도심지 토지 피복 환경에 대하여 분류기법에 따라 Class별로 구분하여 나타내었으며, 각각의 분석결과에 대하여 정확도 검증과 비교를 통하여 도심지 토지피복 분석에 적합한 기법을 산정 하였다. 또한 GIS를 이용하여 토지 분류 기법 별로 분류된 토지피복도와 정확도를 사용자에게 보다 신뢰성 있고 구체적인 정보를 제공할 수 있으리라 판단된다

1.3 연구동향

인공위성 영상의 분광 특성은 지표면의 특정 물체에 대하여 센서가 감응하는 전자기파의 특징을 말한다. 지상의 물체는 그 종류나 상태 등에 따라 전자기 파장의 특성이 틀려진다. 따라서 지표면의 여러 물체의 전자파의 특성을 규명하는 것은 원격탐사(Remote Sensing) 원리의 기본이 되는 것이다. 이러한 원격탐사의 발전은 초기 항공기를 이용한 항공사진 촬영 및 분석기술에 의해 시작되었으며, 1957년 인류 최초의 무인 위성인 Spoutnik가 우주로 발사됨으로써 위성시대의 막이 올랐다. 과거 군사적인 목적으로 활용되던 원격탐사기술이 1990년대 이후부터는 지구자원탐사, 환경 모니터링 및 기상정보 제공 등 인류의 번영을 위한 목적으로 활용되기 시작하였다. 특히 1972년 Landsat 위성이 출현하면서 부터 위성영상 기술의 비약적인 발전을 가져오기 시작했다.

토지피복분류는 원격탐사 자료의 가장 대표적이고 전형적인 응용방법의 하나로, 숲, 초지, 콘크리트 포장과 같은 대지 표면의 물리적 상황을 분류한 것이다. 위성영상을 이용한 토지피복분류에 관련된 연구는 국내외로 꾸준히 연구되어 왔다.

1974년 R.B. Erb는 Landsat 데이터를 이용하여 도시지역에 대한 토지피복을 분류하였으며[4], 1976년 Eugene L. Maxwell과 Barry S. Siegal 등은 다중분광 밴드의 여러 상호관계를 이용한 원격탐사 영상 분류에 관한 연구를 수행하였다[5)6)]. 1980년 Fraiser는 훈련데이터의 크기와 수, 분류방법 등의 중요성을 강조하였다[7]. 1985년 Pain은 호주의 new south wales 동부지역을 대상으로 Landsat 영상을 이용한 토지이용분석의 연구를 하였고[8], 1988년 Stringer 등은 Landsat 영상을 이용하여 지형분석에 관한 연구를 수행하였다[9]. 1990년 Benediksson 등, 1992년 Bischof 등, 1993년 Kamata, 1994년 Solaiman 등은 Landsat TM 데이터의 분류에 신경망을 적용하였다[10)11)12)13)].

1991년 Bolstad와 Lillesand는 위성데이터분류에서 전통적인 기법이 가진 문제는 단지 분광정보에 의존한다고 주장하였고[14], 이러한 문제를 극복하기 위해서 토지피복분류에서 GIS와의 결합을 위한 자동적인 방법을 연구하였다.

피복분류 정확도에 관한 연구도 1990년 Gong과 *H*owarth, 그리고 *J*anssen 등, 1992년 Franklin과 Wilson에 의해서 수행되었다[15)16)17)]. 1996년 Folly 등은 스페인에서 ULSE C-factor를 추정하는 과정에서 두 계절의 Landsat TM 영상과 GIS를 이용하여 피복을 분류하고 지식 기반의 접근 방법을 통하여 토지이용분류의 정확도를 높이고자 하였다[18].

국내의 경우 1986년 연상호는 Landsat MSS 데이터를 이용하여 서울, 대전, 인천 지역의 토지피복상태를 추출하고 분류항목을 도심, 삼림, 수역, 기타의 4개, 5개 항목으로 설정하고 그에 대한 정량적인 분석을 통하여 연구하였다[19]. 1991년 김용일은 원격탐사 데이터의 분류정확도 향상에 관한 연구를 통계적 분류방법을 사용하여 토지피복추출에 적용하였다[20]. 1999년 김홍규는 토지피

복분류에 퍼지이론을 적용하여 토지피복변화를 분석하였고[21], 2001년 전형섭은 2차 계획법을 토지피복분류에 적용하여 토지피복의 Soft 분류 가능성을 분석하였다[22]. 2002년 윤홍식 등은 남한강 유역의 Landsat 영상을 이용하여 영상을 분류하고 토지피복변화량을 분석하였다[23]. 2002년 김헌명은 북한지역의 토지피복분류기법을 제시하였으며[24], 2003년 김태규는 위성영상을 이용하여 정량적 토지피복의 변화탐지에 관한 연구를 수행하였다[25]. 그리고, 2003년 사공호상 등은 고해상도 영상인 IKONOS 영상을 이용하여 토지피복분류 방법에 관한 연구를 하였다[26].

앞에서 언급한 연구동향에서 볼 수 있듯이 1999년을 기점으로 등장한 고해상도 위성영상은 위성영상이 지닌 고유의 특성뿐만 아니라 기존의 항공사진이 갖고 있던 특성까지도 포함하고 있다. 그러므로 세계적으로 이러한 고해상도 위성영상을 이용하여 공간정보를 얻으려는 연구가 활발하게 진행되고 있다.

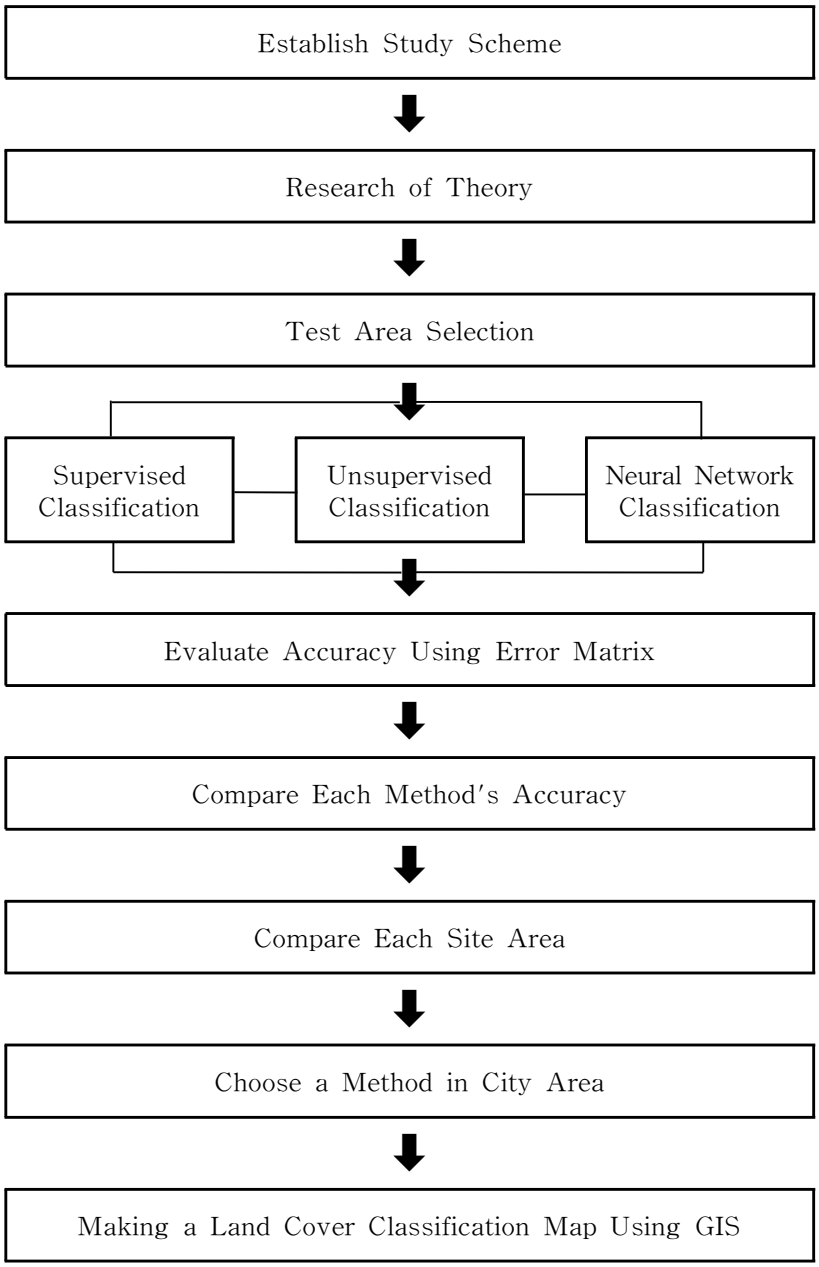
1.2 연구범위 및 방법

본 논문에서는 도심지의 토지피복분류 현황을 기존의 방법보다 더욱 정확히 분석하기 위해서 IKONOS 영상을 이용하여 분석방법에 따른 정확도를 비교·분석하고 GIS를 이용하여 토지피목 현황을 분류기법별로 나타내며, 그 지역의 분류 정확도와 정보를 제시하고자 한다.

영상을 분류하는 방법에는 여러 가지가 있으나, 그 중 가장 효율적인 방법은 분광특성을 이용한 분류방법이다. 건물, 도로, 수목, 논, 밭등 지상의 물체들은 고유의 분광특성을 가지고 있기 때문에 이를 통하여 영상을 분류할 수 있다. 하지만 본 논문의 대상지역은 도심지역이기 때문에 많은 지상 구조물에 의해 서로 다른 분광 클래스(class)등이 복잡하게 나타난다. 따라서 복잡한 분광특성을 나타내는 도심지에 가장 적합한 분류 방법의 산정이 필요하다 연구에서는 먼저 IKONOS 영상에 대하여 1:1,000 수치

지도를 이용하여 도로의 교차점이나 건물의 모서리 등과 같이 영상에서 육안으로 구별할 수 있는 곳을 산정하여 GCP 좌표를 획득하였다. 이렇게 전처리된 영상에서 Test지역을 선정하여 본 연구에서 제시한 영상분석 방법들에 대하여 정확도 검증을 실시하였다. 분류 방법에는 무감독분류(K-means, ISODATA)와 감독분류(Maximum Likelihood Method) 그리고 퍼지 이론을 이용한 신경망 분류를 사용하였다. 총 4개의 방법을 사용하여 대상지역에 대한 분류정확도를 오차행렬을 이용하여 정확도를 산정 하였다.

각각의 분류기법에 대하여 분류된 영상은 Arcview포멧인 shp파일로 변환 하여 각각의 공간데이터와 속성데이터로 구성된 주제도를 제작 하였으며, 그리고 각 항목이미지에 대한 개별적인 정확도도 속성정보에 부여 하였다. 이런 공간정보와 속성정보를 이용하여 최종적으로 GIS를 이용하여 분류기법별 토지피복 분류도를 작성하였다. <Figure 1. 1>은 개략적인 연구 흐름도를 나타내고 있다.



<Figure 1.1> Flow Chart of Research Planning

2. 연구의 이론적 배경

2.1 위성영상 전처리

위성영상 자료는 대기에서 산란과 흡수 등에 의해 또는, 기하학적으로 많은 왜곡이 포함된 자료이기 때문에 실제로 이용하기 위해서는 이러한 부분에 대한 보정이 필요하며, 이러한 과정을 전처리과정이라고 한다. 일반적으로 전처리 과정에는 방사보정과 기하보정으로 크게 나눌 수 있다.

2.1.1 방사보정

방사보정은 원격탐사 자료를 획득한 시점에서 나타나는 센서의 특성, 태양의 고도, 대기조건 및 지형상태 등에 의해서 발생하게 되는 오차를 보정하는 과정이다. 태양의 위치는 계절에 따라 변화하기 때문에 다른 계절에 획득된 자료를 이용하기 위해서는 태양의 위치에 대한 보정이 필요하다. 이러한 보정은 영상자료를 태양 천정각의 코사인값으로 나누어줌으로써 가능하다. 또한 전자기파가 대기를 통과할 때 발생하는 산란에 대한 보정이 필요하며, 이러한 보정은 대기보정이라고 한다. 대기보정은 매우 어려운 과정으로 여러 가지 방법이 이용되지만, 일반적으로 Lowtran 7과 같은 대기보정 모델을 이용하여 보정이 이루어진다.

2.1.2 기하보정

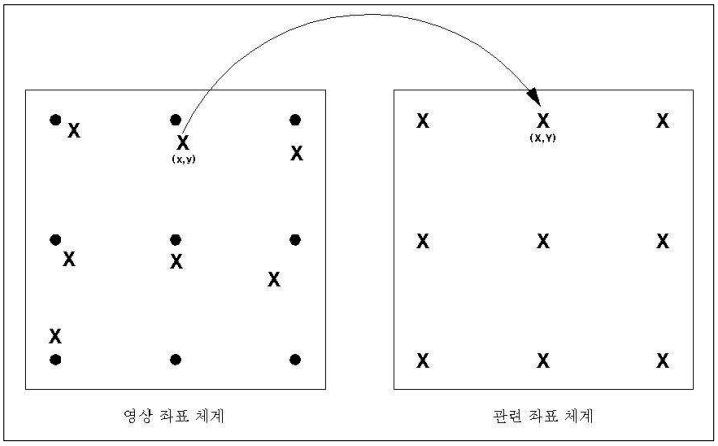
기하보정은 영상내의 좌표값과 실제 좌표값을 일치시키기 위한 과정으로, 최근 원격탐사 자료가 여러 가지 종류의 공간 및 속성자료와 함께 지리정보시스템에서 많이 이용되고 있어 특히 중요한 처리과정이다. 이 과정에서 영상 내에서 지상기준점(GCP : ground control point)를 선정하여 실

제 지형도상의 좌표를 입력하게 되며, 지상 기준점으로는 식별이 가능한 도로의 교차점, 비행장이나 수면과 지표면의 경계 부분이 이용된다. 이러한 지상기준점을 통해 보정 전 영상에서의 좌표값과 보정 후 계산되는 좌표값 사이의 관계식을 만들어 영상 내의 좌표를 변환시키게 된다. 관계식은 일반적으로 식 (2. 1)과 같은 다항식(polynomial)이 이용된다.

$$x = a_0 + a_1 X + a_2 Y + a_3 XY + a_4 X^2 + a_5 Y^2 + \cdots a_k X^m Y^m \qquad (2.\ 1)$$

$$y = b_0 + b_1 X + b_2 Y + b_3 XY + b_4 X^2 + b_5 Y^2 + \cdots b_k X^m Y^m \qquad (2.\ 2)$$

위 식에서 (x, y)는 기하보정 전의 영상좌표이며, (X, Y)는 기하보정 후의 좌표값이다. 실제로 사용되는 모델은 식 (2. 1)의 항에 따라 *H*elmert 변환, Affine 변환, Projection 변환, 이차 등각변환, 다항식 변환 등을 사용한다. <Figure 2.3>은 (x, y)에 상응하는 화소값(X, Y)을 계산하는 기하학적 변환의 모습을 나타내고 있다.



<Figure 2. 1> Geometric Transformations

6개의 미지수를 구하기 위해서는 최소한 3개의 지상기준점이 필요하다.

이러한 지상기준점으로 보통 시간에 따라 변함없는 위치를 가지는 도로의 끝점, 인공구조물, 제방의 끝점 등을 사용한다[27].

좌표변환식의 정확도를 점검하는 방법으로는 식 (2.1)과 식 (2.2)에 의하여 변환된 각각의 지상기준점이 갖게 되는 평균 제곱근오차(RMSE : Root Mean Square Error)를 이용한다. 지상기준점에 의해 재배열(resampling)된 영상은 그 정확도를 점검하여야 하는데 보통 각 지상기준점에 대한 평균 제곱근오차를 계산하는 것이다. 아래 식에서 각 GCP가 가졌던 원래의 좌표값과 변환 후에 갖게 되는 좌표값과의 차이를 통하여 얻어지게 된다.

$$RMSE = \sqrt{(x - x_0)^2 + (y - y_0)^2} \tag{2.3}$$

여기서 x와 y는 변환 후에 갖게 되는 영상의 지상기준점의 좌표값이고, x_0와 y_0는 원 영상의 지상기준점의 좌표이다. 이러한 RMSE는 그 합이 한 화소, 또는 1/2화소보다 작은 값을 갖게 하는 것이 보통이다.

인공위성 영상의 경우, 적절한 전처리과정을 거치고 정확한 지상기준점을 선정하면 Affine 변환과 같은 1차식으로도 충분히 좋은 결과의 보정이 될 수 있으나, 계통적 오차에 의한 왜곡의 보정이 미리 행해지지 않는 경우에는 일반적으로 3차 정도의 다항식이 필요하게 된다. 기하보정을 행한 후 생성된 영상은 픽셀값을 가지고 있지 않게 된다. 즉 각 픽셀이 자신의 위치만을 가지고 있는 영상이 된다. 기하보정의 결과로 나온 최종 영상의 픽셀값은 원 영상의 픽셀값으로부터 추출된다. 이를 픽셀의 재배열(resampling)이라고 하며, 이는 다항식에서 각 계수값들이 최소자승법(least square)등과 같은 수학적인 방법에 의해서 구해지며, 변환식이 결정되면 내삽(interpolation)에 의해서 영상이 재배열 된다. 영상자료의 재배열 방법에는 최근린내삽법(nearest neighbor), 공일차내삽법(bilinear interpolation),

그리고 3차첩입내삽법(cubic convolution) 등이 이용된다.

2.2 영상 분류 이론

영상분류는 원격탐사에서 가장 중요한 분석 기법 중 하나이다. 영상을 구성하는 각각의 화소들 중 비슷한 분광 특성을 갖는 것끼리 집단화시키는 방법이 있는데 이러한 분류 기법은 가장 전통적인 방법이다. 일반적으로 디지털 영상은 지구 표면으로부터 복사되고 방사되는 에너지와 각각 일치하는 요소의 2차원 배열이다. 측정값의 공간적 배열은 영상 또는 영상 공간으로 정의 내려진다. 센서에 의존하여 데이터는 n 밴드내에 저장되며, 디지털 영상 데이터는 일반적으로 8-bit DN 값으로 저장된다. 유사한 DN 값으로 구성된 각각의 집단을 클래스라고 한다. 한 클래스내의 화소는 다른 클래스내의 화소보다는 서로 더 큰 유사성을 지니지만 동일 집단내의 화소라 할지라도 어느 정도 편차는 가지고 있다. 이와 같이 각각의 분광특성에 의해 형성되는 집단을 분광 집단이라고 하며 여러 가지의 분광 집단이 합쳐져서 결국 하나의 정보 집단을 형성하게 되는 것이다. 분광 정보는 대상 지역에 대한 정보가 없는 경우, 물체의 알려진 분광 정보와 취득된 영상 지역의 분광 정보를 비교하여 지역에 대한 환경이나 특성을 대략 적으로 추정할 수 있다. 또한 대상 지역에 대한 지표면의 특성을 이미 알고 있는 경우 지표면의 정밀한 분석이나 변화탐지에 이용할 수 있으며, 파장에 따라 물체의 반사율, 즉 픽셀의 밝기 변화로 영상을 해석하거나 정량적으로 분석을 수행할 수 있다.

이와 같은 영상 분류 기법은 토지 이용도나 토지 피복도, 식생형태, 토양수분함유량 등 여러 가지 분석에서 필수적으로 사용되어지며, 분류 주체에 따라 크게 감독분류와 무 감독분류의 두 가지로 나눌 수 있다. <Table 2. 1>은 여러 가지 분류기법의 특성을 나타내고 있다.

<Table 2. 1> Specific of Classification[27)]

Type \ Specific		Advantage	Disadvantage
Unsupervised	Sequential Clustering	Simple Input Data, Fast	Trial and Error Method
	Statistical Clustering	Acquisition of Exactly Value	Depend upon Input Data
	ISODATA	Minimum Input data	Slow processing speed
	RGB Clustering	Fast	At least 3 Band need
Supervised	Parallelogram Method	Simple and Fast	Incomplete output
	nearest neighbor	Simple Mathematics	Low Accuracy
	Mahalanobis	Clustering's variance Susceptible	Over Classification
	Maximum Likehood	The Most Correct Method	Slow processing speed

2.2.1 감독 분류(Supervised classification)

영상분류의 주요한 과정 중의 하나는 feature space 분할이다. 이는 감독분류에서 Training area를 확인함으로써 클래스들의 분광 특성을 규정하는 사용자의 의해 인지되어진다. 이 방법을 사용하기 위해서 사용자는 항공사진이나 지상 관측자료, 식생도, 지도 등의 자료를 이용하여 자신이 분류하고자 하는 집단의 각 밴드별 화소값을 영상으로부터 추출해야 한다. 이렇게 추출된 영상의 일부분을 표본 집단(training set)이라고 하며, 분류 작업시 표본 집단의 통계적 수치를 계산하여 입력 자료로 사용하게 된다. 분류하고자 하는 집단들 각각에 대한 표본 집단이 완성되면, 이들과 비슷한 통계값을 갖는 화소들을 계속 수집하여 집단의 크기를 확장해 감으로써 전체 영상에 대한 분류 작업을 하게 된다.

이처럼 감독분류는 수행자가 대상 지역을 잘 알고 있어야한다. 작업자는 영상에 의해 포함되어져있는 지역의 클래스를 어디에서 찾아야하는지를 알아야 한다. 작업자에 의해 선택된 클러스터는 주어진 클래스에 대하여 대표적인 데이터를 형성해야 한다. 이는 영상 내 클래스의 다양함이 고려

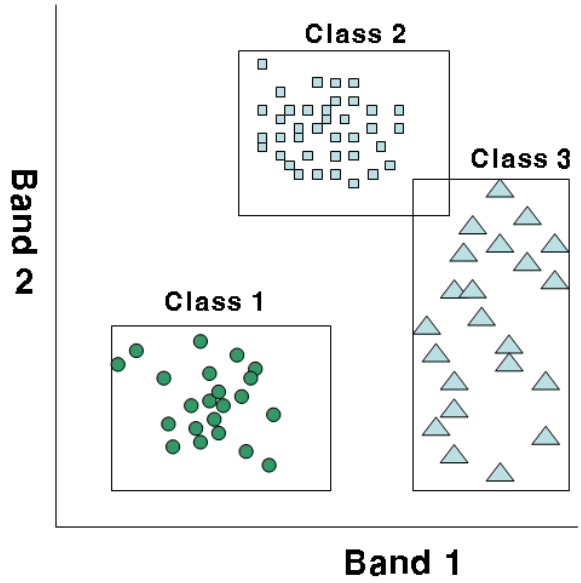
되어져야한다는 것을 의미한다. 또한 다른 클러스터들과 한정되게 중복되지 않게 하여야 하며 만약 그렇지 않다면, 신뢰할만한 분리는 가능하지 않다. 이는 원리적으로 클래스들이 영상분류에 의해 구별 될 수 없다는 것을 의미한다.

이처럼 감독분류는 지역에 대한 사전 지식이 많을수록 적합한 분류 방법이다. 감독 분류 시 단점은 사용자가 지정한 훈련지역에 의하여 클래스를 나누기 때문에 전체 영상이 지니고 있는 정보를 정확히 읽었을 경우에도 기대하지 않은 결과가 나올 수도 있다는 것이다. 대표적인 예는 그림자에 의한 도로 폭의 손실과 기울기에 의한 산림지역의 왜곡 등이 있다.

이와 같은 감독분류의 정보 집단을 결정하기 위해 사용되는 방법에는 여러 가지가 있는데, 그 중 평행육면체분류(parallel-piped classifier), 최대우도분류(maximum likelihood class), 최소거리분류(maximum distance classfifier) 등이 일반적으로 많이 사용되고 있다[28].

2.2.1.1 평행육면체분류

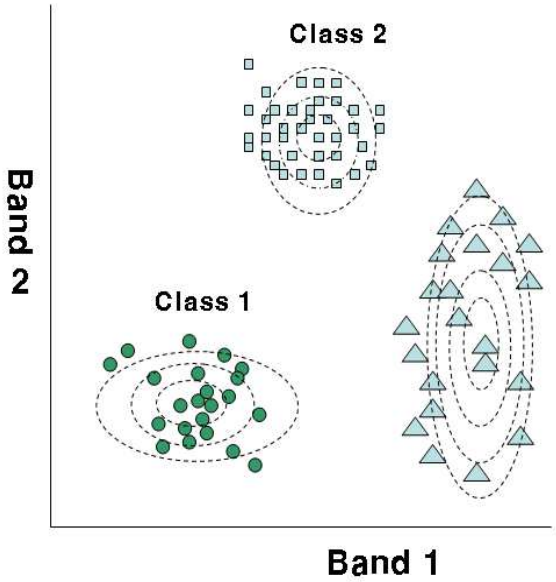
평행육면체분류는 화상의 화소 값들을 통계적인 상한과 하한의 테두리 속에서 분류하는 방법이다. 이때 한계선 결정은 각 항목의 추정평균값에 표본 표준편차를 더하여 상한과 하한으로 결정하거나, 분석자가 경험에 의존하여 임의로 결정할 수 있다. 이렇게 결정된 한계선은 2차원 분류에서는 사각형, 3차원 분류에서는 육면체가 되어, 그 형태 안에 존재하는 화소는 모두 한 항목으로 분류되며 어디에도 속하지 않는 경우는 미분류 항목으로 분류된다. 이 방법에서는 각 밴드에 대해 항상 일정한 한계치와 화소값이 비교되므로 빠르고 간단한 장점이 있으나, 한계구역이 중복되거나 이를 벗어나는 화소들을 분류하기 위한 판별력이 부족하다.



<Figure 2.2> Parallel-Piped Classifier

2.2.1.2 최대우도 분류

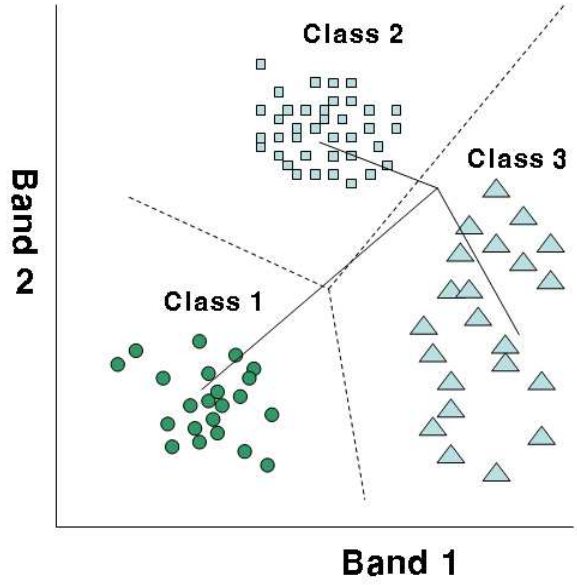
최대우도분류는 현재 가장 널리 사용되는 방법으로 주어진 화소값이 각각의 Training site에 포함될 확률에 기초하여 분류 항목을 설정하는 방법이다. 화상의 각 화소가 특정 분류 항목으로 분류되는 것은 앞의 최단거리 분류처럼 분류항목의 중심으로부터의 거리로부터 결정되는 것이 아니라 패턴공간의 공간적 분포에서 공분산과 관계되는 등확률값 분포선의 차원과 방향성에 달려 있다. 이 방법은 밴드의 히스토그램이 정규분포를 이룬다는 가정하에 계산이 수행된다. 공분산 행렬을 이용함으로써 분류항목별 분광 공간에서서의 분포의 다양성을 고려한 것이 장점이지만, 전산처리시 많은 계산량을 필요로 한다. 또한 밴드의 히스토그램이 정규분포를 이룬다는 가정은 자료가 정규분포를 이룰 경우 높은 분류 정확도를 보이지만, 레이더 자료처럼 자료의 히스토그램이 정규분포를 이루지 않을 경우 심각한 분류 오차를 보일 수 있다.



<Figure 2.3> Maximum Likelihood

2.2.1.3 최소거리 분류

최소거리분류는 판별하고자 하는 화소값을 각 집단에 속하는 훈련집단 화소의 값과 그 사이의 거리를 비교하여 가장 가까운 훈련집단 화소가 속하는 집단에 화소를 활당하는 방법이다. 이 방법은 비교적 빠르고 간단한 계산을 통해 분류가 수행되는 장점이 있으나, 어느 항목에도 속할 수 없는 미분류되어야 할 화소도 분류가 되므로 정확도 면에서 문제가 있다. 또 한 도시지역처럼 분산이 큰 자료 분포 항목에서 실제로는 포함되어야 하지만, 평균에서 멀리 떨어져 있다는 이유로 다른 항목으로 분류되는 경우가 발생할 수 있다.



<Figure 2.4> Minimum Distance Classfication

2.2.2 무감독 분류(Unsupervised classification)

무감독 분류는 대상 지역에 대해서 현지 조사가 어렵거나 파장별 반사 특성이 유사하여 감독 분류를 수행하기 힘든 경우에 적용되는 분류기법이다. 군집화(clustering)라고도 불리우며, 분류항복에 대한 표본집단과 같은 사용자의 사전 설정없이 화상을 구성하는 화소값들의 통계적 특성만을 이용하여 분류를 수행여 몇가지 집단으로 분류한다. 분석자는 무감독 알고리즘을 실행하기 전에 군집의 수, 군집간의 최소거리, 군집간의 한계거리 등을 지정하고 이 초기변수들에만 의존하여 컴퓨터가 군집화 분류를 수행하게 된다. 이때 군집은 분류항목의 내용과는 상관없이 분광특성만을 나타내는 분광집단을 나타내므로 실행 후 분석자가 항목의 정의, 즉 정보집단의 정의를 내려주어야 한다. 그러나 이렇게 구성된 각각의 집단은 초기 기준점을 임의로 위치시켰기 때문에 우리가 원하는 것과 같이 정확한 분류가 이루어졌다고 볼 수 없다.

지상의 많은 지역은 정치적 혹은 지형적 이유 때문에 접근이 불가능하

므로 많은 경우 감독분류에 필요한 Training site의 설정이 매우 어렵거나 불가능한 경우가 있다. 감독 분류의 경우 분석지역내의 분류 항목 수의 선택에 따라 분석결과에 상당한 차이가 있지만 시간에 따라 지표면 현상이 변하므로 정확한 분류 항목의 수를 결정하는 것 또한 매우 어려운 경우가 많다. 반면 무감독 분류는 대상 지역의 특성에 대한 사전 정보 없이 지역적 특성과 관련된 통계적 특성 값의 추정을 통한 분류를 수행할 수 있는 장점이 있다. 그러나 무감독 분류는 각 집단의 통계값이 반복 횟수에 따라 순차적으로 변화하기 때문에 정보 집단과 분광 집단 사이의 관계가 애매한 경우가 생길 수 있기 때문에 대상물의 형상, 분포특성 등을 고려하여 대상물을 구분해야 하는 단점이 있다

이를 수행하는 대표적인 알고리즘으로는 K-mean 클러스터링과 ISODATA등이 있다. K-mean clustering 방법은 정해진 개수의 초기 군집에 대한 중심벡터를 임의로 설정하여 군집화를 수행하는 방법이다. 초기에 임의로 각 집단의 중심이 선택되면, 각각의 화소에서 중심까지의 거리를 계산하여 최단거리를 갖는 집단을 찾아내어 군집화를 수행하게 된다. 초기 중심을 임의로 설정하므로 이에 의한 영향을 최소화하기 위하여 전체 영상에 대한 군집화를 반복적으로 수행한다.

ISODATA clustering 방법은 Iterative Self Organizing Data Analysis Technique의 약자로 반복 계산시 각 군집의 새로운 평균이 계산되며 군집 정리시 군집의 표준편차값을 분석하여 편차값이 너무 큰 경우 두 개의 군집으로 분리할 수 있다. 이러한 군집 분리 기능을 가짐으로써 읽어들인 모든 화소는 분광거리가 가장 가까운 군집으로 할당된다.

2.2.3 신경망 분류(Neural Network classification)

신경망은 뉴런(Neuron)이라 불리는 연결인자를 상호 병렬적으로 연결하여 구축되는 비모수, 비선형기법의 일종으로 분류(Classification)에 사용될 수 있다. 그 예는 "Neural network approaches versus statistical methods in classification of multisource remote sensing data(*J.* A. Benediktsson, P. *H.* Swain. and O. K. Ersoy,*J*uly 1990)"에서 볼 수 있다. 신경망의 구조는 입력층, 은닉층, 출력층으로 단순하지만 상호 연결된 처리 요소인 수많은 뉴런들로 구성된 정보처리 구조들이다. 입력층은 입력 정보에 대하여 각 입력값에 가중치(weight)를 곱해서 은닉층으로 전달한다. 은닉층은 입력층으로부터 각기 다른 가중치에 의해 변환된 자료들을 모두 합하여 전달함수를 통하여 출력하게 되고 이 값들은 다시 은닉층과 출력층 사이의 가중치와 곱하져 출력층으로 보내진다. 출력층은 은닉층으로부터 각기 다른 가중치에 의해 변환된 자료들을 받으며, 각 신경단위는 합산된 자료들을 출력값의 목적에 따라 binary값의 출력을 위한 Signum함수, 연속값의 출력을 위한 선형 또는 비선형 함수에 의하여 변형시켜 출력한다. 역전파 알고리즘(back propagation algorithm)은 다층 피셉트론에서 입력패턴과 출력패턴에 대하여 오차를 줄여가는 과정에 적용되는 방법으로 이는 정보처리 시스템의 목적에 맞도록 합리적으로 각 층에서 입출력 패턴의 가중치를 조절하는 것이다.

입력층 I에 입력자료가 입력되었을 때 출력층 O로 출력되는 네트워크를 생각해보자. 신경망에서 모든 연결은 연결강도(weight)라 불리는 수치값을 가진다.(w_{ij}: 단위 i에서 j의 연결의 강도를 나타낸다.) 뉴런i는 다른 뉴런들의 출력치 O_j로부터 그리고 연결강도w_{ij}로부터 순입력(net_i)을 계산한다. 대부분의 모델에서 편기값(bias, b_j)로 불려지는 수치값이 순입력에 더해진다. 이 값에 적용된 함수 f(활성함수)는 뉴런의 출력치 O_i를 생성한다. 식

으로 이것은 다음과 같이 표현된다.

① 입력자료 : $X_i, i = 1, \cdots, N_l$ (2.4)

(N_l = 입력 자료의 차수)

② 은닉층 Node로의 가중된 입력 : $h_j = \sum_{j=0}^{N_l} w_{ij} X_i + \theta_j,\ j = 1, \cdots, N_H$ (2.5)

(N_H = 은닉층 Node의 수, θ_j = 은닉층 j번째 Node의 Bias입력,

w_{ij} = 입력층 I번째 Node와 은닉층 j번째 Node사이의 가중치)

③ 은닉층의 Node로부터의 출력 : $H_j = f(h_j)$ (2.6)

$f(u)$(*Sigmoid* 함수) = $\dfrac{1}{1 + e_{(-u)}}$ (2.7)

$f(u)$ (*Hyperbolic Tangent* 함수) = tanh(u) =$\dfrac{1 - e_{(-u)}}{1 + e_{(-u)}}$ (2.8)

④ 출력층 Node로의 가중된 입력 : $h_k = \sum_{j=0}^{N_w} w_{ij} X_i + \theta_k,\ k = 1, \cdots, N_o$ (2.9)

(N_o = 출력층 Node의 수, θ_k = 출력층 k번째 Node의 Bias 입력,

w_{ij} = 은닉층의 j번째 Node와 출력층 k번째 Node 사이의 가중치)

⑤ 출력층 Node로의 출력 : $Y_k = f(h_k)$ (2.10)

$f(u)$(선형 변환 함수) = $f(u)$ (2.11)

$f(u)$(Signum 변환 함수)= (1 if u > 0, 0 if u = 0, -1 if u > 0) (2.12)

위와 같은 학습과정은 각 입력값에 대해 출력을 계산하는 순반향처리와 실제값과 출력값 사이에서 발생하는 오차를 역으로 진행시켜 이 오차를 최소화하도록 층과 층 사이의 가중치를 변화, 조절하는 역방향처리로 구성된다. 가중치의 초기치는 -0.1과 0.1사이의 난수를 사용한다. 순반향처리는 입력 자료가 신경회로망을 통하여 출력 자료를 계산하는 과정이 입력층, 은닉층, 출력층을 차례로 통과하여 출력값을 산출하는 과정이다. 역

방향처리과정은 출력층에서의 실제값과 출력값과의 오차를 최소화하도록 가중치를 조절하고 은닉층의 출력값을 얻어 은닉층과 입력층 사이의 가중치를 반복적으로 조절하는 형태를 갖는다[29].

2.3 분류영상의 정확도 검증기법

원격탐사를 통하여 취득된 주제도나 이와 관련된 통계치들을 유용하게 사용하려면 데이터의 분류정확도를 정량적으로 명확히 평가할 수 있는 방법이 있어야 한다. 분류정확도 평가는 원격탐사의 다른 연구분야에 비하여 뒤늦게 연구된 분야로 아직까지도 많은 연구에서 간단한 수치로서 분류정확도를 단순히 평가하고 있다. 하지만, 이러한 분류정확도 평가는 위치정확도를 완전히 무시하는 것으로서 항목의 전체 양이 위치에 관계없이 평가된다. 이러한 위치정확도를 포함하지 않는 정확도 평가는 매우 높은 정확도를 산출하지만 전체적인 오차가 지역적으로 평형을 이루지 않을 때, 심각한 오류를 범하게 된다. 따라서 분류정확도 평가를 정확하게 수행하기 위해서는 분류결과가 참조데이터와 비교되어야 하며 <Table 2.2>와 같은 오차행렬로 평가된다.

<Table 2.2> Error Matrix

	A	B	C	D	E	F	Row Total
A	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	$X_{1.}$
B	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}	X_{26}	$X_{2.}$
C	X_{31}	X_{32}	X_{33}	X_{34}	X_{35}	X_{36}	$X_{3.}$
D	X_{41}	X_{42}	X_{43}	X_{44}	X_{45}	X_{46}	$X_{4.}$
E	X_{51}	X_{52}	X_{53}	X_{54}	X_{55}	X_{56}	$X_{5.}$
F	X_{61}	X_{62}	X_{63}	X_{64}	X_{65}	X_{66}	$X_{6.}$
Column Total	$X_{.1}$	$X_{.2}$	$X_{.3}$	$X_{.4}$	$X_{.5}$	$X_{.6}$	N

오차행렬은 정방행렬로서 분류항목의 수만큼의 열과 행으로 구성되는데 각각 참조데이터(reference data)와 원격탐사데이터로부터 생성된 분류결과를 나타낸다. X_{ij}에서 i=j인 대각 행렬 원소들의 픽셀들을 정확하게 분류된 결과들이며 i≠j인 행렬 원소들은 분류오류를 나타낸다. 또한 이 행렬에서 $X_{.j}$ (j=1, 2, ..., 6)는 열에 대한 합을 영상분류에서 각 분류등급들의 픽셀수를 나타내고, X.i (i=1, 2, ..., 6)는 행에 대한 합으로 실제 연구대상지역에 대해 각 분류등급들의 픽셀수가 되며 N은 전체 픽셀수가 된다.

또한 오차행렬은 포괄적 오차(commission error)와 배타적 오차(omission error)를 표현하면서 각 분류항목의 정확도를 명확히 묘사할 수 있기 때문에 정확도를 표현하는데 매우 효과적인 방법이다. 그러나 문제는 원격탐사 분류결과와 비교하기 위하여 편의되지 않은 지상참조정보를 어떻게 수집할 것인가이다. 따라서 불편의 지상참조정보를 수집하기 위해서는 항목당 수입되어야할 샘플수, 샘플링 방법 그리고 다변량통계와 같은 적절한 결과 묘사방법 등이 고려되어야 한다.

이러한 오차행렬을 사용하여 필요한 정확도를 구할 수 있는데, 이를 통

해 구해지는 정확도는 크게 전체분류에 대한 정확도와 각 분류등급들에 대한 정확도를 의미한다. 정확도를 구하는 방법은 전체적인 정확도와 Kappa 계수를 구하는 방법이 있다. 전체정확도는 정확하게 분류된 픽셀수들을 검정지역전체에 대한 픽셀수로 나누어서 구해지며 정확하게 분류된 결과들만을 사용하여 정확도를 나타내므로 매우 대략적인 정확도가 된다. 이에 비하여 Kappa 계수값은 정확하게 분류된 결과뿐만 아니라 분류오류들도 고려하여 정확도를 계산하기 때문에 전체정확도보다 훨씬 정확한 정확도를 나타내는 기준이 된다.

전체정확도에서 각 항목의 정확도를 계산하는 것은 분석자가 항목의 올바르게 분류된 픽셀수를 해당 열이나 행의 픽셀 총수로 나누어야하기 때문에 매우 복잡하다. 일반적으로 항목의 올바르게 분류된 총 픽셀수는 참조데이터로부터 유도된 그 항목의 총픽셀수(즉, 열의 총수)에 의해 나누어진다. 이러한 통계는 올바르게 분류될 참조데이터의 확률을 나타내며, 이것을 배타적 오차라 한다. 분석가는 어떤 지역이 얼마나 잘 분류될 수 있는가에 관심이 있기 때문에 이러한 통계를 생산자 정확도(producer's accuracy)라 하기도 한다. 그리고 한 항목에서 올바르게 분류된 화수의 총수가 그 항목에서 실제적으로 분류된 픽셀의 총수에 의해 나누어진 것을 포괄적 오차라 한다. 사용자 정확도(user's accuracy) 혹은 신뢰도라 불리는 이 측정치는 분류된 지도에서 한 픽셀과 지상의 그 항목을 나타낼 확률이다[31].

$$A_p = \frac{X_{ii}}{X_{+i}} \qquad (2.13)$$

(여기서, A_p는 생산자 정확도)

$$A_u = \frac{X_{ii}}{X_{i+}} \qquad (2.14)$$

(여기서, A_u는 생산자 정확도)

위의 식(2.4) 및 (2.5)에서 X_{i+}는 오차행렬 i번째 행의 픽셀의 합, X_{+i}는 오차행렬에서 i번째 열의 픽셀의 합을 의미한다.

Kappa 분석은 일치도(measure of agreement) 혹은 정확도인 K_{hat} 통계치(Kappa 추정치)를 산출하며 다음 식(2.6)과 같이 계산된다[30].

$$Kappa = \frac{N\sum_{i=1}^{n} x_{ii} - \sum_{i=1}^{n}(x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{n}(x_{i+} \times x_{+i})} \qquad (2.15)$$

여기서 N은 관찰된 픽셀의 총 개수, n은 오차행렬에서 행의 수, x_{ii}는 i행과 I열번째의 관찰된 개수, x_{i+}와 x_{+i}는 i행과 i열번째의 한계 총계를 의미한다. 이를 정확도로 표현하면 다음과 같이 나타낼 수 있다.

$$Kappa = \frac{(1/N)\sum_{i=1}^{n} x_{ii} - (1/N^2)\sum_{i=1}^{n}(x_{i+} \times x_{+i})}{1 - (1/N^2)\sum_{i=1}^{n}(x_{i+} \times x_{+i})} = \frac{P_0 - P_e}{1 - P_e} \qquad (2.15)$$

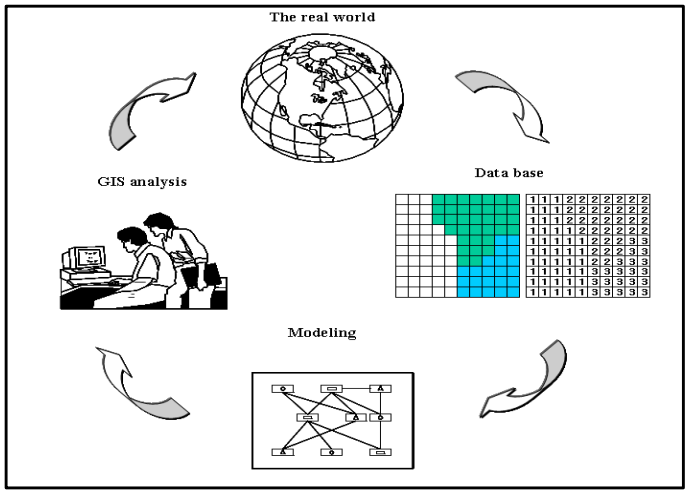
여기서 P_0는 일치된 픽셀 개수, 즉 전체 정확도를 의미하며, P_e는 일어날 수 있는 일치된 픽셀 개수를 의미한다. 이렇게 정의되는 Kappa 계수는 통계적인 의미에서 전체 정밀도로부터 한계분포를 이용하여 계산되는 우연의 일치에 의한 효과를 제거한 결과로 나타나게 되며, 분산행렬로 표현되는 분류의 결과가 완전히 우연으로 인한 결과일 경우에는 0의 값을 가지게 된다. 양의 Kappa 계수는 기대된 일치보다 클 경우에만 가능하다. 실제 일치성이 1에 접근하고, 가능 일치성이 0에 접근하면 K는 1로 접근한다. 이러한 경우는 아주 이상적인 경우로 볼 수 있으며, 실제로 Kappa

계수의 범위는 통상적으로 0~1사이에 있다. 전체정밀도의 경우에는 단순히 행렬의 대각선 요소들만을 이용하여 계산하게 되므로 주변분포, 즉 누락오차를 고려할 수 없으나, Kappa 계수의 경우는 행렬의 주변분포를 이용하여 계산되어지는 값이므로 행렬의 대각선 요소가 아닌 요소들도 고려하게 되며, 따라서 이러한 내부오차 및 내부오차의 영향에 대한 정보를 포함하게 된다. 그러므로 우연으로 인한 결과를 포함하고 있는 전체정확도보다는 Kappa 계수의 경우가 분류의 정확도를 표현하는 방법으로 적당하다고 볼 수 있다[31].

2.3 GIS 데이터 모델

GIS(Geographic Information System)는 지리적 자료를 수집, 저장, 분석, 출력할 수 있는 컴퓨터 응용시스템으로 지형공간에 관한 모든 정보를 컴퓨터에 저장, 이를 바탕으로 각종 계획 수립과 의사결정 및 산업활동을 효율적으로 지원할 수 있도록 만든 첨단 정보시스템을 말한다.

실세계의 공간은 무한하고 매우 복잡하므로 이러한 공간에 가장 근접하기 위해서는 다양한 정보와 체계적인 자료들을 필요로 한다. 지형공간정보체계에서의 가장 큰 목적은 이러한 자료들을 이용하여 사용자가 목적하고자하는 세계를 모델링함으로써 접근하고자 하는 것이다. 따라서 데이터 모델이란 사용자의 목적에 따른 여건을 적합화시켜서 가장 근접한 형태의 목적에 따른 환경으로 추구하는 것이다. 이러한 일련의 과정을 표현하면 <Figure 2.5>과 같다.



<Figure 2.5> GIS Data Model Notion

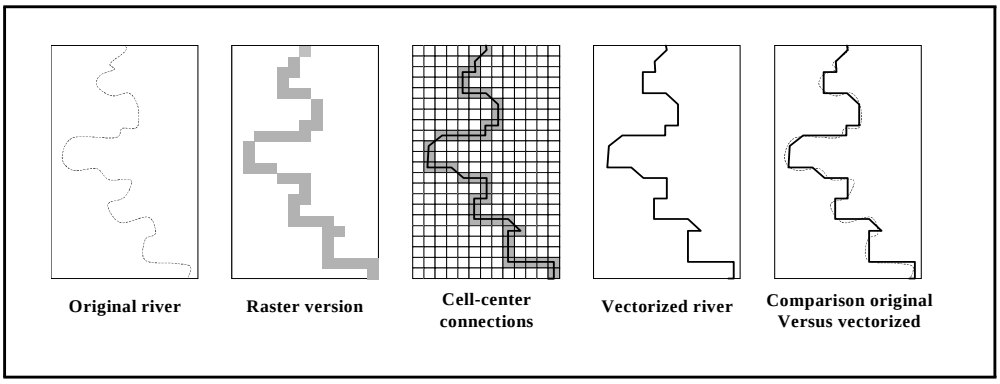
2.3.1 공간데이터베이스

공간데이터베이스는 GIS에서 사용되는 도형과 속성자료를 합친 모든 정보를 입력하여 보관하는 정보의 저장소이다. 수치지도의 형태로 공간데이터베이스에 저장된 각종 정보는 정보의 검색과 수정 및 보완, 분석을 위한 추출에 용이하도록 도형자료와 속성자료가 서로 효율적으로 연결되어야 한다. GIS에서 사용되는 데이터베이스자료는 공간자료(Spatial Data)와 속성자료(Attribute Data)로 구분할 수 있다. 공간자료 혹은 도형자료(Graphic Data)는 도면성에서의 위치를 나타내는 공간객체에 관한 자료를 말한다. 일반적으로 지표상에 존재하는 물체들을 공간객체(Spatial Objects)라 하며 이들 공간객체는 점(point), 선(line, arc), 폴리곤(polygon) 또는 면(area)으로 구분된다. 공간자료는 Vector 구조와 Raster 구조로 나눌 수 있다.

Vector 자료는 도형을 구성하는 점, 선, 면의 위치와 길이 및 면적을 정확하게 표현하기 위한 방식이다. 이 자료는 뛰어난 위상관계를 구축할 수 있기 때문에 하천, 도로, 지적, 등고선, 경계선 등의 표현에 용이하고, 위치와 속성의 일반화가 가능하며, 정확한 그래픽 표현이 가능하다. 그러나 자

료구조가 복잡하고 지도중첩이 어려운 단점도 있다.

Raster 자료는 좌표에 의해 위치가 결정된 셀(cell)들을 이용하여 포인트, 라인, 폴리곤의 공간데이터를 표현한다. 모든 셀은 하나의 특성을 가지면서 독립적으로 숫자들에 의해 저장된다. 이 자료의 특징은 데이터 파일이 matrix형태로 구성되므로 각 grid의 공간적 위치는 matrix내의 row와 column에 의해 자동적으로 표시된다. 따라서 자료구조가 간단하고, 지도중첩이나 위성자료와의 결합이 용이하며, 단위별로 위상관계가 동일하나, 네트워크가 어렵고 그래픽 자료의 양이 방대한 점도 있다. <Figure 2.6>은 래스터구조를 벡터화, 벡터구조를 래스터화로 상호 전환하는 개념을 표현하고 있다[32].



<Figure 2.6> Conversation of Data from Raster to Vector Format

3. 분광특성을 이용한 위성영상 분류

3. 1 연구대상지역

본 논문에서 사용된 위성영상은 미국의 Space Imaging사에서 발사한 IKONOS 위성에서 촬영된 영상이다. IKONOS 위성은 민간용 위성이 수행할 수 있는 최상의 공간해상력을 지니고 있으며, 높은 공간 해상력 외에도 11비트로 양자화(quantisation)를 하기 때문에 복사 해상도가 높다. IKONOS영상의 특성은 <table 3.1>과 같다.

<Table 3. 1> Specific of IKONOS OSA

Name of Satellite	IKONOS
Orbit	Satellite Altitude: 680km, Sun Cycle: Am 10:30hr, Pass Cycle: 14Day
Sensor	OSA(Optical Sensor Assembly)
Swath width	11km(12㎛ CCD elements)
Off-track-viewing	±50° All Direction
Revisit time	1～3Day
spectral band, ㎛	0.45～0.52(1), 0.52～0.60(2), 0.63～0.69(3), 0.76～0.90(4), 0.45～0.90(PAN)
Ground pixel size	1m(PAN), 4m(bands 1～4)

대상지역은 부산광역시의 동래구 일원 2.654×3.051의 면적으로 분광특성으로 토지 피복을 분류하기 위하여 기하보정과 방사보정이 실시된 4m해상도의 다중분광(RGB)영상만을 이용하였다. 본 연구지역은 도심지와 산림지역만 있는 곳으로 도심지의 토지피복의 분류에 있어 적합한 영상이라고 판단된다.

분류시 지표의 공간 객체의 명칭을 부여하는 것과 이들 정보를 몇 개의

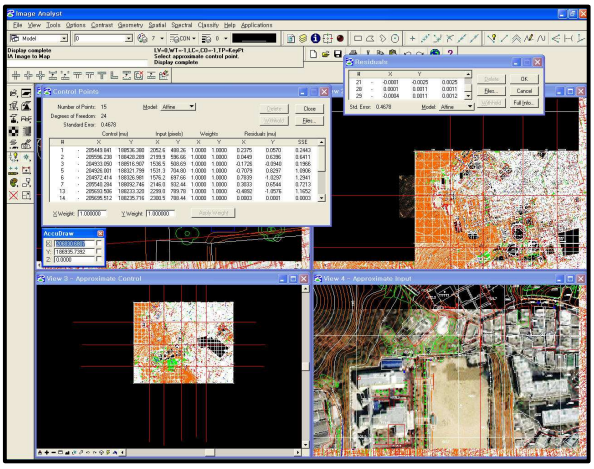
클래스로 분류할 것인지를 결정하여야 한다. 본 연구에서는 미국지질조사국(United States Geological Survey : USGS)의 토지이용/토지피복 분류체계의 Level Ⅱ의 분류기준을 고려하여 분류하였다. <Figure 3.1>은 연구 대상지역을 나타낸 IKONOS영상이다.

<Figure 3. 1> IKONOS Image of Study Area

3. 2 위성영상의 전처리

본 연구에서는 국립지리원에서 발간한 1:1,000 수치지도를 이용하여 기하보정을 실시하였다. 사용된 소프트웨어는 Z/I Imaging사의 Image Analyst를 이용하여 분석하였다. 먼저 대상지역에 15개의 GCP(Ground Control Point)를 도로 교차점이나 건물의 모서리 부분과 같이 영상에서 판독이 용이한 곳을 선정하였으며, 오차가 크게 발생하는 지역에 대해서 삭제하는 방식으로 8개의 지상기준점을 선정하였다. Image to map 모드에서 Affine 변환을 사용하여 기하 보정을 수행 하였으며, 기하보정을 행한

후 생성된 영상은 픽셀값을 가지고 있지 않으므로 각 픽셀은 자신의 위치만을 가지고 있는 영상이 되므로 원 영상의 픽셀값으로부터 최종영상의 픽셀값을 추출하는 영상재배열방법이 필요하다. 본 연구에서는 영상재배열방법 중 최근린내삽법(nearest neighbor)을 이용하여 영상을 재배열 하였다. <Figure 3. 2>는 기하보정을 실시하는 모습을 나타내고 있으며, 잔차는 X축으로 0.3425, Y축으로 0.5453, RMSE는 0.6692로 나타났다.



<Figure 3. 2> Geometric Correction

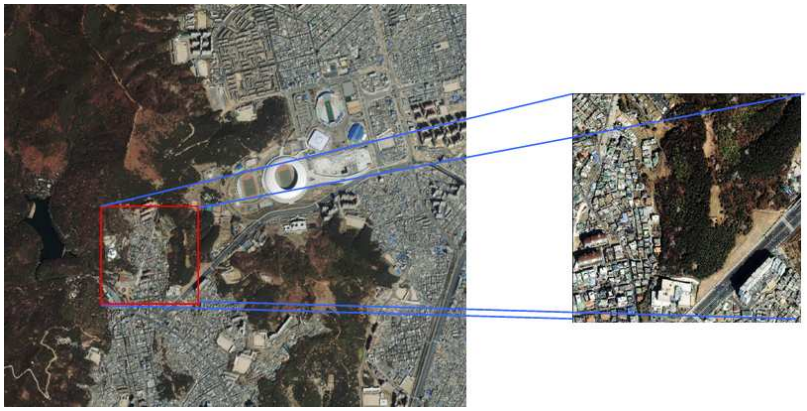
<Table 3. 2> Control Point of RMSE

Point ID	Control		Input[Pixels]		Residuals		SSE
	X	Y	X	Y	X	Y	
1	205449.841	188536.380	2052.6	488.26	0.2375	0.0570	0.2443
2	205596.238	188428.289	2199.9	596.66	0.0449	0.6396	0.6411
3	204933.050	188516.907	1536.5	508.69	-0.1726	-0.0940	0.1966
4	204926.001	188321.799	1531.3	704.80	-0.7079	0.8297	1.0906
5	204972.414	188326.981	1576.2	697.66	0.7839	-1.0297	1.2941
6	205540.284	188092.746	2146.0	932.44	0.3033	0.6544	0.7213
7	205693.506	188233.320	2299.0	789.78	-0.4892	-1.0576	1.1652
8	205695.512	188235.716	2300.5	788.44	0.0003	0.0001	0.0003
RMSE	0.4883						

3. 3 Test 지역 선정

본 논문에서 Test지역은 전형적으로 도시의 특성을 가진 지역을 선정하였다. Test 지역의 선정은 먼저 분류항목이 모두 포함되어 있는 지역을 선정하는 것이 중요하다. 따라서 Test 지역은 500×500 Pixel의 크기로 산림, 건물, 도로, 초지, 나지의 지역이 고르게 분포하고 있는 지역을 선정하였고 <Figure 3. 3>에 나타내었다.

감독분류는 분류하고자 하는 대상물에 대한 사전지식이나 알고리즘을 알고 있어야 훈련지역(training site)를 설정할 수 있다. 이에 반해 무감독분류는 보다 자동화된 방법으로, 영상의 수치 값만으로 분류를 수행함으로 훈련지역을 만들 수는 없지만 판단 기준의 guid line을 설정해 줄 수 있다. 감독분류에 있어서 분류하고자 하는 항목의 수와 사용한 데이터에 의해 분류 정확도가 영향을 받는다. 본 연구에서는 1:1,000, 1:5,000의 지형도와 현지조사 및 항공영상의 육안분석을 통하여 각 분류항목의 훈련지역을 선택하였다. 정확도 평가는 error matrix을 작성하여 수행하였다.



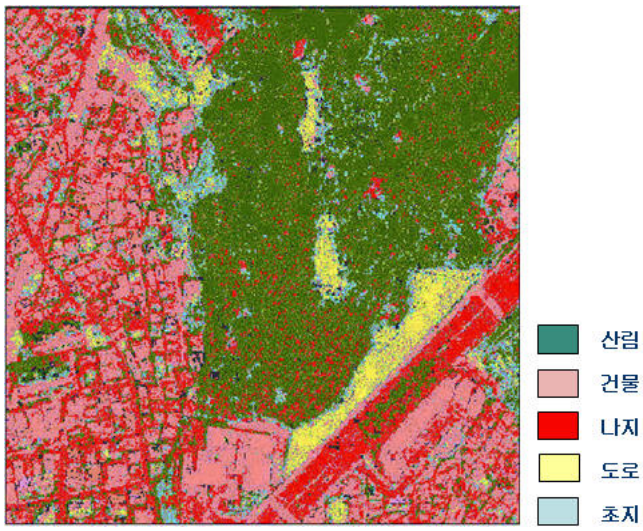
<Figure 3. 3> Test Area

3. 4 감독 분류

본 연구에서는 Region Growing 표본추출 기법과 분류결과에 확률 통계적인 의미를 갖는 등 우수한 특징을 가지고 있는 최대우도법을 이용하여 감독 트레이닝을 실시하였다. 트레이닝의 결과는 군집이나 훈련지역(training site)를 의미하는 signature의 집합으로 나타난다. 개개의 signature는 하나의 class에 대한 site이고 하나의 class에 화소들을 할당하기 위한 결정규칙과 함께 적용된다.

훈련데이터의 추출은 수치지도와 현지답사를 통하여 화면상에서 이루어졌으며, 항공사진에서 취득된 영상에 대하여 토지피복형태가 변형되지 않은 지역을 선별하여 동일한 지리적 위치의 훈련데이터를 추출하였다. 설정된 분류 항목에 대하여 각각 10×10Pixel 규모로 지역적으로 고르게 선정하였으며 분류된 영상은 LFP 필터를 적용하여 분류시 발생한 노이즈나 미분류 지역을 제거 하였다.

분류 결과는 <Figure 3. 4>에서 보는 바와 같이, 건물, 산림, 초지, 나지, 도로로 구분하였다. 훈련지역의 데이터에 대한 밴드별 평균값과 표준편차를 살펴봄으로써 분류 항목간의 분류정확도를 판단할 수 있으며, 단변량 통계값과 공분산 값은 <Table 3. 2>,<Table 3. 3>과 같으며, 분류 정확도는 <Table 3. 4>과 같이 오차 행렬로 나타내었다..



<Figure 3. 4> Land Cover Test Area Using Supervised Classification

<Table 3. 3> Covariance Matrix(Supervised Classification)

Layer	Band 1(Red)			Band 2(gre)			Band 3(Blu)		
	Band1	Band2	Band3	Band1	Band2	Band3	Band1	Band2	Band3
Road	805.561	728.600	586.202	728.600	687.590	555.310	586.202	555.310	480.129
Building	1309.136	1006.031	843.525	1006.031	1006.881	1000.763	843.525	1000.763	1162.626
Tree	438.813	273.039	219.702	273.039	237.042	161.703	219.702	161.703	130.651
Grass	531.724	479.828	349.537	479.828	489.814	364.799	349.537	364.799	309.167
Bare	221.402	180.434	212.315	180.434	168.264	206.318	212.315	206.318	276.650

<Table 3. 4> Univariate Matrix(Supervised Classification)

Layer	band 1				band 2				band 3			
	Min	Max	Mean	Std. Dev	Min	Max	Mean	Std. Dev	Min	Max	Mean	Std. Dev
Road	3	239	81.64	28.38	5	232	86.08	26.22	8	226	84.41	21.91
Building	37	222	150.24	36.18	59	220	147.16	31.73	43	216	135.91	34.10
Tree	7	120	51.83	15.40	12	96	51.83	15.40	10	82	43.80	11.43
Grass	65	169	124.41	23.06	58	149	109.70	22.13	48	125	87.14	17.58
Bare	114	180	154.14	14.88	95	161	135.67	12.97	55	139	107.17	16.63

<Table 3. 5> Error Matrix(Supervised Classification)

Class	Tree	Bare	Grass	Road	Building	Producer's Accuracy (%)	User's Accyracy (%)
Tree	55	1	0	2	0	91.04	96.82
Bare	0	4	0	0	1	70.00	87.5
Grass	0	2	6	0	1	78.78	70.00
Road	4	0	0	10	5	92.86	48.15
Building	2	0	1	1	86	90.76	97.30

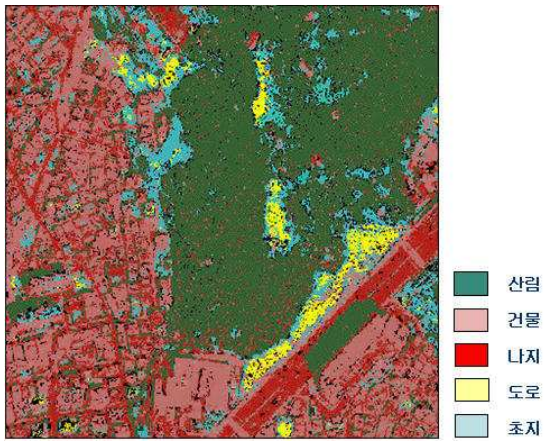
3. 5 무감독 분류

본 연구에서는 무감독 분류 방법 중 K-means Training와 ISODATA를 이용하였다. 이 방법은 K-means, ISODATA Classifier를 이용하여 영상 분류를 수행하는데 여기서 Classifier는 영상 자료의 밝기값에 있어 파장별 거리를 이용하여 각 요소들에 대해 유한개의 집단을 구성하는 방법을 이용하여 분류를 수행한다. 초기에는 임의로 각 집단의 중심이 선택되어지며, 미지의 화소에서 이 중심까지의 거리를 계산하여 가장 짧은 거리를 갖는 화소들을 찾아 집단을 구성하게 된다. 일단 모든 화소에 대해서

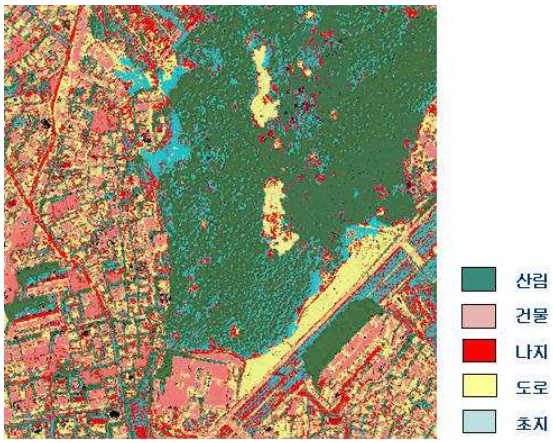
소집한 집단이 결정되면, 이들로부터 새로운 집단 중심이 계산되어지며, 모든 집단에 대해 더 이상 중심값의 변화가 없을 때까지 이러한 과정을 반복하여 수행하였다.

먼저 ISODATA 분류는 200번 반복 계산을 하였으며, 다소 시간이 오래 걸리는 단점이 있지만 정확한 분류가 가능하였다. 그리고 분류 항목이 5개인 점을 감안하여 Lumping distance는 20.00의 중간 값을 지정하였다. 이렇게 분류 항목은 총 15개로 이는 다시 참조자료와 일치도가 높은 5개의 항목으로 재분류를 실시하였다.

K-means 방법은 ISODATA와 같이 반복 횟수를 200번으로 지정하고 convergence distance는 1.5로 지정하였다. 이는 clustering에 있어 각 픽셀들 간의 분광특성에 대한 거리를 나타내는 것으로 본 논문에서는 평균값을 적용하였다. <Figure 3. 5>, <Figure 3. 6>는 K-means와 ISODATA를 이용하여 테스트 지역의 토지 피복을 분류한 영상을 나타내고 있으며, 또한 각각의 공분산 값, 단변량분 통계값의 분류 정확도는 <Table 3. 5> ~ <Table 3. 10>와 같이 나타났다. 분류 결과 건물과 나무의 분류 정확도가 가장 높게 산정되었으며, 반면 초지와 나지의 사용자 정확도와 생산자 정확도는 낮게 나타났다. 이는 감독 분류와 마찬가지로 영상취득 시기에 따른 분광값이 비슷한 class간의 교차 분류되었기 때문이라고 판단된다. 무감독 분류된 영상 또한 LFP 필터를 적용하여 분류 시 발생 한 노이즈를 제거 하였다.



<Figure 3. 5> Land Cover of Test Area sing K-means



<Figure 3. 6> Land Cover of Test Area Using ISODATA

<Table 3. 6> Covariance Matrix(K-means)

Layer	Band 1(Red)			Band 2(gre)			Band 3(Blu)		
	Band1	Band2	Band3	Band1	Band2	Band3	Band1	Band2	Band3
Road	226.38	96.07	16.50	96.07	146.18	111.19	16.50	111.19	178.73
Building	312.29	278.96	289.85	278.96	289.51	320.82	289.85	320.82	392.70
Tree	69.28	55.90	41.81	55.90	57.82	44.73	41.81	44.73	47.12
Grass	112.45	54.31	29.71	54.31	72.71	52.09	29.71	52.09	73.39
Bare	249.24	134.71	64.31	134.71	147.06	140.53	64.31	140.53	225.71

<Table 3. 7> Univariate analysis Matrix(K-means)

Layer	band 1				band 2				band 3			
	Min	Max	Mean	Std. Dev	Min	Max	Mean	Std. Dev	Min	Max	Mean	Std. Dev
Road	65	169	124.41	23.06	58	149	109.70	22.13	48	125	87.14	17.58
Building	124	255	177.61	17.67	141	255	174.31	17.02	99	255	160	19.82
Tree	0	79	32.62	8.32	0	58	33.02	7.60	0	75	32.12	.56.86
Grass	14	116	57	10.60	24	93	55.28	8.53	0	130	50.09	8.57
Bare	53	205	134.72	15.79	90	167	130.60	12.13	23	173	114.95	15.02

<Table 3. 8> Error Matrix(K-means)

Class	Tree	Bare	Grass	Road	Building	Producers Accuracy (%)	User's Accyracy (%)
Tree	57	2	1	3	5	93.44	84.00
Bare	0	5	0	0	0	71.42	100
Grass	1	0	5	0	8	71.42	35.71
Road	3	0	1	10	9	76.92	43.48
Building	0	7	0	0	86	89.26	100

<Table 3. 9> Covariance Matrix(ISODATA)

Layer	Band 1(Red)			Band 2(gre)			Band 3(Blu)		
	Band1	Band2	Band3	Band1	Band2	Band3	Band1	Band2	Band3
Road	134.88	74.76	45.05	74.76	101.31	84.76	45.05	84.76	115.74
Building	312.31	273.12	264.77	273.12	299.50	329.58	264.77	329.58	422.95
Tree	118.01	86.65	68.37	86.65	89.21	74.96	68.37	74.96	85.01
Grass	223.35	92.15	8.54	92.15	134.93	189.97	8.54	109.97	178.79
Bare	222.56	104.57	32.43	104.57	124.92	118.60	32.43	118.60	192.07

<Table 3. 10> Univariate analysis Matrix(ISODATA)

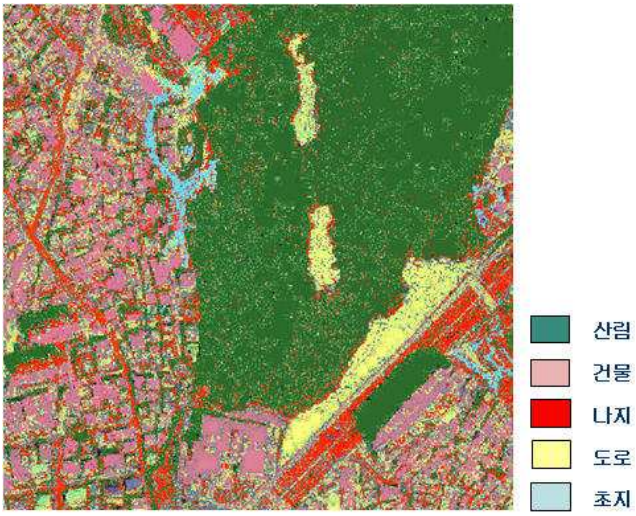
Layer	band 1				band 2				band 3			
	Min	Max	Mean	Std. Dev	Min	Max	Mean	Std. Dev	Min	Max	Mean	Std. Dev
Road	23	132	71.55	11.61	37	106	66.96	11.61	69	102	125.87	13.86
Building	129	255	184.22	17.67	149	255	181.93	17.31	114	255	170.14	20.57
Tree	0	69	40.42	10.86	0	69	37.98	9.44	0	82	34.30	9.22
Grass	39	171	107.91	14.94	65	139	103.27	11.62	28	168	92.88	11.62
Bare	76	207	144.01	14.92	96	175	139.50	11.18	69	182	125.87	13.86

<Table 3. 11> Error Matrix(ISODATA)

Class	Tree	Bare	Grass	Road	Building	Producers Accuracy (%)	User's Accyracy (%)
Tree	54	0	1	2	6	88.52	85.71
Bare	0	5	0	0	0	71.42	100
Grass	0	2	5	0	3	71.42	50.0
Road	4	0	0	10	4	76.92	55.56
Building	3	0	1	1	95	87.968	95.0

3. 6 신경망 분류

K-means, ISODATA 방법과 감독분류 방법의 Algorithm들은 단순히 영상 자료에 나타나는 각 Pixel들을 그 밝기의 통계적 거리 관계에 따라 비슷한 Pixel들끼리 Grouping을 해주는 일종의 Clustering 방법을 사용한다. 하지만 각 Pixel들 사이의 통계값을 이용한 Clustering 방법은 단순히 수학적 계산에 의해 각각의 Class가 결정되어 지므로, 각 Class의 경계 부분이나 Shade 효과가 큰 산지 등에 존재하는 Pixel들에 대해서는 정확한 영상분류가 불가능하다. 이러한 단점을 보안하기 위해서 Neural Network 이론을 적용한 분류 방법을 이용하였다. 이 방법은 Puzzy 이론을 이용하여 컴퓨터의 영상 분류 작업에 인공지능적인 요소를 가미한 것으로써 분류하기에 애매한 Pixel 값에 대해서도 정확한 분류를 가능하게 할 수 있다. 분석시 계산의 반복 횟수는 200번으로 지정 하였고 총 class수는 5개로 구분하였다. <Figure 3. 7>는 신경망을 이용하여 분류된 영상을 나타내고 있으며, <Table 3. 11>, <Table 3. 12>, <Table 3. 13>은 공분산값, 단변량분 통계값, 분류 정확도를 나타내고 있다.



<Figure 3. 7> Land Cover of Test Area
Using Neutral Network classification

<Table 3. 12> Covariance Matrix(Neural Network)

Layer	Band 1(Red)			Band 2(gre)			Band 3(Blu)		
	Band1	Band2	Band3	Band1	Band2	Band3	Band1	Band2	Band3
Road	230.57	108.88	37.18	100.00	157.38	122.80	37.10	122.80	183.87
Building	203.96	180.87	187.09	180.07	177.25	193.85	187.09	193.85	235.68
Tree	209.01	160.76	121.62	160.76	154.02	119.42	121.62	119.42	113.54
Grass	181.35	72.25	-8.99	72.25	181.76	90.43	-8.99	98.43	187.36
Bare	128.01	61.26	23.67	61.26	87.15	91.50	23.67	91.50	150.61

<Table 3. 13> Univariate analysis Matrix(Neural Network)

Layer	band 1				band 2				band 3			
	Min	Max	Mean	Std. Dev	Min	Max	Mean	Std. Dev	Min	Max	Mean	Std. Dev
Road	22	159	88.23	15.18	49	137	83.97	12.55	10	168	74.73	13.56
Building	143	255	189.71	14.28	163	255	186.75	13.31	121	255	176.87	15.35
Tree	0	116	43.08	14.46	0	83	42.10	12.41	0	112	39.39	18.66
Grass	53	191	126.66	13.47	90	154	123.04	10.09	23	172	107.44	13.69
Bare	105	212	156.77	11.31	126	183	152.00	9.34	69	179	137.38	12.27

<Table 3. 14> Error Matrix(Neural Network)

Class	Tree	Bare	Grass	Road	Building	Producers Accuracy (%)	User's Accyracy (%)
Tree	57	0	1	3	3	93.44	89.06
Bare	0	4	1	0	5	57.14	40.0
Grass	1	2	5	0	0	71.42	62.5
Road	3	1	0	10	5	76.92	52.63
Building	0	0	0	0	95	87.96	100

3. 7 분류 결과 분석

분류기법에 따라 각각의 정확도는 Error Matrix를 이용하여 산정하였으며 총 196개의 참조점을 이용하였다. 먼저 최대우도법을 이용한 감독분류의 정확도를 평가한 결과 전체 정확도는 89.80%, Kappa 계수는 0.83으로 나타났다. 분류 항목 중 산림지역의 정확도가 가장 높게 나왔으며, Ortiz등이 제시한 Kappa 값에 따라 최상의 분류 정확도로 평가된다. 생산자 정확도는 최대 값은 90.16%이고 사용자 정확도는 96.19%이다. 나지와 초지의 경우가 정확도가 낮게 나타났다. 이는 영상의 취득 시기가 겨울인 관계로 나지와 초지가 서로 교차되어 분류되었다고 판단 된다.

무감독 분류 방식인 K-means와 ISODATA의 전체 정확도는 86.22%와 83.16%로 ISODATA 방법의 정확도가 높게 나타났으며, Kappa 값은 0.73, 0.77으로 산정되었다. 이 방법들은 분류 시간이 짧은 장점이 있었으나 참조 자료를 구축하기 힘든 단점이 있었다. 또한 반복 작업을 얼마나 할 것인가, 일치도를 어느 정도로 설정할 것인가 등에 결과가 많이 좌우 되었고, 참조자료에 따라 정확도가 많은 영향을 받았다. 감독분류는 도심지역과 같이 반사특성이 다양한 지역의 훈련 데이터를 산정하는데 있어 어려운 점이 있었으며, 이러한 훈련데이터 결과에 의해 정확도가 많이 좌우되었다. 값은 Ortize et al.(1997)에 의해 설정된 분류 평가에 대한 Kapp계수 범위에 있어 분류 정확도가 우수한 범위를 나타내었다. 그리고 Neutral Network 무감독 분류의 정확도는 87.24%, kappa 계수는 0.79로 가장 높은 분류 정확도를 나타내었다. 또한 감독분류와 무감독분류에 있어 초지와 나지의 분류 정확도가 낮게 나타난 반면 도로와 산림, 건물의 분류 정확도는 상대적으로 높게 나타났다. 이는 영상의 취득 시기가 겨울인 관계로 초지 지역과 나지 지역의 분광특성이 비슷하여 서로 교차 분류되었다고 예상된다. 또한 각 분류 기법 별로 단변량 통계치와 공분산 값을 산정 하여 분류 항목에 대한 기보적인 통계자료를 구축하여 분광특성에 대한 분류정도를 알 수 있었다. <Table 3.14>는 분류 방법에 있어 정확도와 Kappa 값을 나타내고 있으며, <Table 3.15>는 Ortiz등이 제시한 Kappa 값에 대한 분류 정도를 나타내고 있다.

<Table 3. 15> Classification of Overall Accuracy, kappa

Classification Method	Overall Accuracy(%)	Kappa
Maximum Likelihood	89.80	0.83
ISODATA	86.22	0.77
K-means	83.16	0.73
Neutral Network	87.24	0.79

<Table 3. 16> Kappa Quality

Kappa	Quality
< 0.00	Worst
0.00～0.20	Poor
0.20～0.40	Proper
0.40～0.60	Good
0.60～0.80	Best
0.80～1.00	Excellent

<Table 3. 17> Number of Class Pixel

Method / Class	Maximum Likelihood	K-means	ISODATA	Neural Network
Road	1205423	1368984	1483975	1287091
Tree	3456114	3059057	3055574	3459479
Building	1823861	1905529	1767898	1686613
Grass	732203	784166	918984	753484
Bare	879753	884678	870923	910687

최대 우도법의 감독분류와 K-means, ISODATA, Neutral 무감독 분류의 정확도를 비교한 결과 최대 우도법을 적용한 감독분류의 정확도가 가장 높은 값을 나타내었다. 그러나 최대 우도법을 적용한 감독분류는 도시지역과 같은 다양한 반사특성을 보이는 경우, 훈련데이터를 선정하기 힘든 단점이 있었으나 대상 지역의 정보가 사전에 많이 구축되어 있어 정확도가 가장 높게 나타났다고 판단된다.

무감독분류에 있어 ISODATA, K-means 분류는 Clustering 방법 각 Pixel들 사이의 통계값을 단순한 수학적 계산에 의해 각각의 Class를 결정하여, Class의 경계 부분이나 Shade 효과가 큰 산지 등에 존재하는 Pixel들에 대해서는 정확한 분석이 불가능 하였다. 반면, Neutral Network 방법은 Puzzy 이론을 사용하여 분류하기 애매한 Pixel 값에 대해서도 분류가 가능하기 때문에 다른 분류 방법보다 정확도가 높게 나타났다. 그리고 <Table 3.16>과 같이 분류 항목별 픽셀의 수에 따라 면적 분포를 계산할 수 있었으며, 평균 도심지는 22.17%, 산지 40.23%, 초지 9.8%, 나지 10.46%로 나타났으며, 이는 Shp파일로 변환시켜 면적 계산이 가능하다.

각각의 분류 정확도와 Kappa 계수를 참조하여 대상지역 전체를 각각의 분류 방법으로 토지피복 분류를 실시하여 대상지역의 토지피복도를 GIS를 이용하여 작성 하였다.

4. GIS를 이용한 토지피복도 작성

4. 1 토지피복도 작성 방법

대상지역의 토지피복 분류도를 작성하기 위해서 RS기법으로 분류된 각 항목이미지(Raster Data)를 ArcView포맷인 S*H*P파일(Vector Data)로 변환하여, 이미 알고 있는 오류를 수정한 후, ArcView를 이용하여 공간데이터와 속성데이터로 구성된 GIS Data로 저장한 후에 토지피복 분류도를 제작하였다. 먼저 분류된 토지이용도에 원하는 분류항목을 선택하면 그 항목에 대한 면적, 전체 정확도, Kappa 계수의 정보를 화면상에 출력 할 수 있도록 GIS를 이용하여 Database를 구축하였다. <Figure 4.1>은 GIS를 이용하여 토지피복도를 작성하는 방법을 나타내고 있다.

Classified Result Image: (Land-Cover Image)

⬇

Conversation of Data from Raster to Vector Format

⬇

Adjustment of Error

⬇

Allowance of Spatial and Attribute Data

⬇

Construction of DEM Target Area by Digital Map

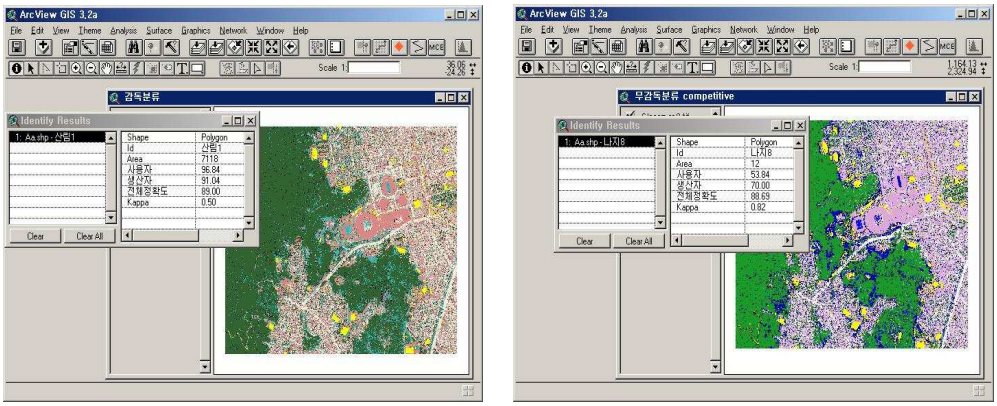
⬇

Land-Cover Map

<Figure 4. 1> Flow Chart of Land-Cover Map by GIS

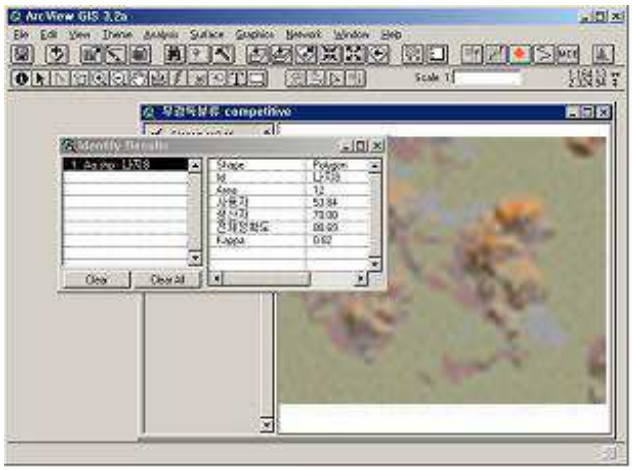
4. 2 토지피복도 작성

GIS Data 상에 분류 정확도와 속성정보를 제공함으로써 사용자가 요구하는 방법에 대한 분류 결과를 좀더 신뢰성 있고 체계적으로 제공하도록 데이터베이스를 이용하여 토지피복도를 작성하였다. 또한 다른 분류 방법과 비교·판단하여 지역에 따른 적합한 분류 방법을 선택할 수 있으며, 지역에 따른 토지피복에 분류 방법에 대한 통계값으로 활용하여 향후 분류 방법의 선택에 있어 중요한 판단 자료 활용될 수 있다고 판단된다. <Figure 4. 2>, <Figure 4. 3>는 분류 방법에 대한 대상지역의 토지피복 분류도에 대한 각각의 정확도와 면적, Kappa계수를 나타내고 있다.

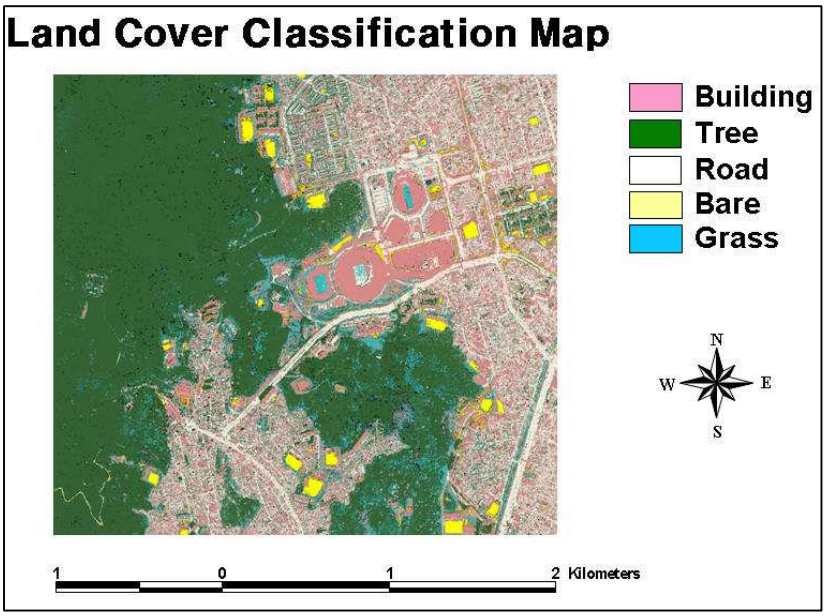


<Figure 4. 2> Allowance of Spatial and Attribute Data on GIS

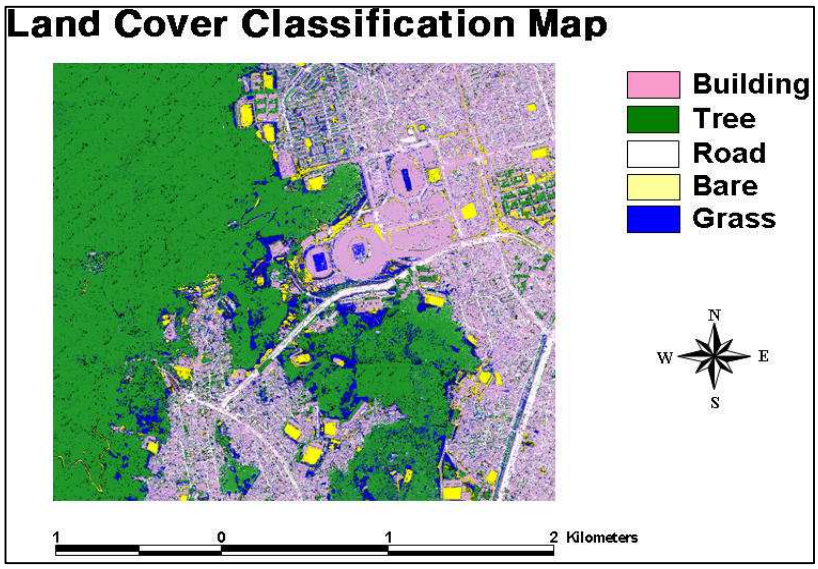
또한 수치지도를 이용하여 각 분류 항목에 대한 DEM을 추가적으로 속성정보를 제공 하였다. 먼저 수치지도에서 등고선과 표고점을 추출하여 Shp 파일로 변화시켜 Tin을 이용하여 DEM을 나타내었다. 이런 분류지역의 DEM은 사용자로 하여금 토지 이용 현황과 더불어 공간적인 정보를 제공함으로써 보다 구체적인 공간 정보를 제공할 수 있다. 그리고 최종적인 토지피복 분류도는 <Figure 4. 5>,<Figure 4. 6>과 같다.



<Figure 4. 3> DEM on Land Use Map



<Figure 4. 4> Land Cover Classification Map of Research of Site(Supervised Classification)



<Figure 4. 5> Land Cover Classification Map of Research of Site(Neural Network)

5. 결 론

IKONOS위성영상을 이용하여 도심지역 토지 피복 분류를 영상 분류 기법별로 분석하고, 분석한 결과와 GIS를 이용하여 토지피복 분류도를 작성한 결과 다음과 같은 결론을 얻었다.

1. 분광특성을 이용한 분류방법 중 최대우도법을 이용한 감독 분류가 전체 정확도는 89.80%, Kappa 계수는 0.83으로 4가지 분류 방법중 가장 높은 정확도를 나타내었다.

2. 주관성을 배제한 분류 방법에 있어 Neural network를 적용한 분류 기법이 전체정확도 87.24%, Kappa 계수 0.79으로 가장 높은 정확도를 나타내었다. 무감독 분류 시 Class의 경계나 산지 등의 Shade 영향으로 교차 되는 Pixel값에 대하여 Neural network적용으로 보다 정확한 분류가 가능 하였다.

3. 고해상도 인공위성영상으로 도심지의 피복분류를 실시한 결과 최대우도법을 적용한 방법이 정확도면에서 적합하며, 분류시 시간적인 면에서는 Neural Network분류 방법이 도심지 피복 분류에 있어 적합하다고 판단된다.

4. 향후 고해상도 위성영상을 이용하여 보다 세분류 항목에 있어 분광특성을 이용한 분류 방법의 정확도 분석이 필요하며, 분류 시스템과 GIS의 상호 연동 체계를 개발하여 보다 정확하고 신속한 자료 갱신의 연구가 필요하다고 판단된다.

참고문헌

1. 조재현, GIS와 위성영상을 이용한 충주시 도시변화량 분석, 석사학위논문, 충주대학교, 2004.
2. 김윤수, 도시분석을 위한 원격탐사자료 활용 방안 연구, 석사학위논문, 계명대학교, 2004.
3. 사공호상, IKONOS 영상의 토지피복분류 방법에 관한 실증 연구, 한국지리정보학회, Vol.6, No.3, pp.107~116, 2003.
4. R.B. Erb, ERTS-1 Urban land use analysis, ERTS-1 investigation(ER-600), *NASA TMX-58121*, Vol. 5, Period July 1972~June 1973.
5. Eugene L. Maxwell, Multivariate system analysis of Multispectral imagery, *PE&RS*, Vol. 42, No. 9, pp. 1173~1186, 1976.
6. Barry S. Siegal, Michael *J.* Abrams, Evaluation of clasiification techniques applied to geologic mapping using LANDSAT data, *42nd Annual Meeting of ASP Technical Papers*, 1976.
7. Fraiser, B.E., *H*.F. Shovic, Statistical methods for determining land use change with aerial photographs, *PE&RS,* Vol. 46, No. 8, pp. 1067~1077, 1980.
8. Pain, C. F., Mapping of landforms from landsat imagery : an example form eastern new south wales, Australia, Remote Sensing of Environment, Vol. 17, No. 1, pp. 55~65, 1985.
9. Stringer, W. *J.*, Groves, *J.* E., and Olmsted, C., Landsat determined geographic change, PE&RS, Vol. 54, No. 3, pp. 347~351, 1988.
10. Benediktsson, *J*.A., Swain, P.*H*., Ersoy, O.K., Neural network approaches versus statistical methods in classification of

multisource remote sensing data, *IEEE TGRS*, Vol. 28, pp. 846~552, 1990.

11. *H.* Bischof, W. Schneider, A.*J.* Pinz, Multispectral classification of LANDSAT images using neural networks, *IEEE TGRS*, Vol. 30, No. 3, pp. 482～490, 1992.
12. Sei-ichiro Kamata, Eiji Kawaguchi, Application of neural network approach to classify multi-temporal LANDSAT images, *IGRASS93*, pp. 716～718, 1993.
13. B. Solaiman, M.C. Mouchot, A comparative study of conventional and neural network classification of multispectral data, *IGRASS94*, 1994.
14. Bolstad, Lillesand, Automated GIS integration in land cover classification, *ASPRA/ACSM Annual Convention*, Vol. 3, 1991.
15. Gong, P. and P.*J.* *H*owarth, The use of structural information of improving land-cover classification accuracies at the rural-urban fringe. *PE&RS*, Vol. 56, No. 1, pp. 67～73, 1990.
16. *J*anssen, L.L.F., M.N. *J*aarssma and E.T.M. van der Linden, Integration topographic data with remote sensing for land-cover classification, *PE&RS*, Vol. 56, No. 11, pp. 1503～1506, 1990.
17. Franklin, S.E. and B.A. Wilson, A three-stage classifier for remote sensing of mountain environment, PE&RS, Vol. 58, No. 4, pp. 449～454, 1992.
18. Folly, A., M.C. Bronsveld, and M. Clavaux, A knowledge-based approach for C-factor mapping in Spain using Landsat TM and GIS, Int. *J.* Remote Sensing, Vol. 17, No. 12, pp. 1401～1415, 1996.
19. 연상호, 원격탐사기법에 의한 서울, 대전, 인천지역 토지이용 분류 연구, 대한원격탐사학회지, Vol. 2, No. 2, pp. 69～77, 1986.

20. 김용일, 인공위성 원격탐사 데이터의 분석정확도 향상에 관한 연구, 박사학위논문, 서울대학교, 1991.
21. 김홍규, 퍼지 집합이론을 이용한 원격탐사 영상의 최대우도 토지피복 분류 및 변화탐지 기법의 개발, 박사학위논문, 강원대학교, 1999.
22. 전형섭, 원격탐사 데이터의 분류 정확도 향상에 관한 연구, 박사학위논문, 전북대학교, 2001.
23. 윤홍식, 조재명, 안영준, 인공위성 영상을 이용한 남한강 유역의 토지피복 변화량 검출, *한국측량학회지*, Vol. 20, No. 2, pp. 41～49, 2002.
24. 김헌명, 다중영상 합성을 통한 북한 지역의 토지피복 분류기법 제시에 관한 연구, 석사학위논문, 인하대학교, 2002.
25. 김태규, 위성영상을 이용한 정량적 토지피복변화탐지에 관한 연구, 석사학위논문, 전북대학교, 2003.
26. 사공호상, 임정호, IKONOS 영상을 이용한 불투수지표면 분석방법에 관한 연구, *한국지리정보학회지*, Vol. 6, No. 3, pp. 107～116, 2003.
27. Robert A. Schowengerdt, Techniques for image processing and classification in remote sensing, pp.99～109
28. 김태규, 위성영상을 이용한 정량적 토지피복변화탐지에 관한 연구, 석사학위논문, 전북대학교, 2003.
29. 김홍태, 신경망을 이용한 인공위성영상의 토지이용도 분류 정확도 비 교, 대한토목학회지, Vol, 1999, No.4, pp 515～518.
30. 박주영, 토지피복분류 정확도 향상을 위한 위성합성 방법 비교 연구, 석사학위논문, 서울시립대학교, 2003.
31. 텍스쳐 분석에 의한 고해상도 위성영상의 분류, 석사학위논문, 상명대학교, 2004.
32. 이희연, GIS, 지리정보학, 법무사, 2003

감사의 글

먼저 배움에 있어 부족함을 채워주시고 지혜를 주신 하나님께 영광을 돌립니다. 지난 2년간의 시간들은 저의 삶에 있어 많은 지식과 경험을 쌓을 수 있는 중요한 시간으로 평생을 살면서 잊지 못할 추억이 될 것입니다.

많이 부족한 저에게 학문과 배움의 자세 그리고 겸손한 마음을 지니도록 가르쳐주신 이종출 지도교수님게 진심으로 감사드리며, 언제 어디서든 지도 교수님의 말씀과 가르침을 깊이 새기며 생활하겠습니다.

그리고 바쁘신 와중에도 논문을 완성할 수 있도록 지도해주신 이환우 교수님, 정두회 교수님께 감사들 드리며, 학부 및 대학원 생활동안 학문을 가르쳐 주신 김종수 교수님, 손인식 교수님, 장희석 교수님, 이종섭 교수님, 이동욱 교수님, 이영대 교수님, 김명식 교수님, 정진호 교수님, 국승규 교수님, 이상호 교수님께도 감사의 마음을 전합니다.

또한 연구실 발전을 위하여 항상 노력하시고 이끌어 주신 서동주 박사님, 논문을 완성하는 동안 언제나 깊은 배려와 격려를 해주신 노태호 박사님, 멀리에서 근무하셔도 항상 관심과 격려를 해주신 장호식 박사님, 연구실 실장을 맡아 노력해주신 김진수 선배님, 학과 일과 함께 저에게 많은 조언을 해주신 조용재 선배님, 연구실 생활하는 동안 언제나 곁에서 힘이 되어준 연구실 동기인 성호 형과 , 윤성에게 감사하며, 연구실의 많은 선배님들을 모시고 열심히 생활해준 유영화 후배님, 서재수 후배님, 남일현 후배님에게 감사의 마음을 전합니다.

특히, 2년동안 생활하면서 힘들고 어려울 때 마다 힘이 되어준 친구 덕배, 재혁, 광래, 지현, 대용이게 고마움을 전하며, 어려운 여건에도 나의 선택을 받아들여, 항상 곁에서 따뜻하게 격려해주며 과정을 마칠 수 있도록 헌신적으로 도와준 소영이에게 진심으로 감사의 마음을 전합니다.

또한 언제나 든든한 후원자 역활을 해준 누나에게 감사를 드립니다.

끝으로 지금까지 못난 자식을 변함없는 사랑으로 지켜봐 주신 부모님께 하해와 같은 감사의 마음을 전하며 이 논문을 바칩니다.