연구논문

분광 유사도 커널을 이용한 하이퍼스펙트럴 영상의

Support Vector Machine(SVM) 분류

Support Vector Machine Classification of Hyperspectral Image using Spectral Similarity Kernel

최재완* · 변영기** · 김용일*** · 유기윤**** Choi, Jae Wan · Byun, Young Gi · Kim, Yong Il · Yu, Ki Yun

要 旨

통계학습이론에 기반하고 있는 Support Vector Machine(SVM)은 구조적 위험 최소화원리를 바탕으로 하는 학습 알고리즘이다. 일반적으로 SVM은 비선형 경계를 결정하고 자료를 분류하기 위해서 커널(kernel)을 사용한다. 그러 나 기존의 커널들은 두 벡터간의 내적이나 거리차를 이용하여 유사도를 측정하기 때문에 하이퍼스펙트럴 영상분류 에 효과적으로 적용될 수 없다. 본 논문에서는 이를 해결하기 위해서 분광유사도커널(Spectral similarity kernel)을 제안한다. 분광유사도 커널은 두 벡터의 거리차와 각 차이를 모두 계산하는 지역적 커널로 하이퍼스펙트럴 영상의 분광특성을 효과적으로 고려할 수 있다. 이를 검증하기 위해서 Hyperion 영상에 polynomial kernel, RBF kernel을 사용한 SVM 분류기와 분광유사도 커널을 사용한 SVM 분류기를 적용하여 토지피복분류를 시행하였다. 분류결과를 통해서 분광유사도 커널을 사용한 SVM 분류기가 정량적, 공간적으로 가장 우수한 결과를 보임을 확인하였다. **핵심용어 :** Support Vector Machine(SVM), 분광유사도 커널, 토지피복분류, 하이퍼스펙트럴

Abstract

Support Vector Machine (SVM) which has roots in a statistical learning theory is a training algorithm based on structural risk minimization. Generally, SVM algorithm uses the kernel for determining a linearly non-separable boundary and classifying the data. But, classical kernels can not apply to effectively the hyperspectral image classification because it measures similarity using vector's dot-product or euclidian distance. So, This paper proposes the spectral similarity kernel to solve this problem. The spectral similarity kernel that calculate both vector's euclidian and angle distance is a local kernel, it can effectively consider a reflectance property of hyperspectral image. For validating our algorithm, SVM which used polynomial kernel, RBF kernel and proposed kernel was applied to land cover classification in Hyperion image. It appears that SVM classifier using spectral similarity kernel has the most outstanding result in qualitative and spatial estimation.

Keywords : Support Vector Machine (SVM), spectral similarity kernel, land cover classification, hyperspectral.

1. 서 론

원격탐사 자료는 지구표면에 대한 다양한 정보를 포함 하고 있기 때문에 환경, 군사, 지질 등의 많은 분야에 다양 하게 활용되고 있으며 최근 타 응용기술의 발달로 원격탐 사 자료의 활용범위는 점차 증대되고 있다. 특히 분광 폭 이 좁은 많은 수의 연속적인 밴드를 가지고 있는 하이퍼 스펙트럴 영상을 이용한 원격탐사에 대한 관심이 높아짐 에 따라 하이퍼스펙트럴 영상의 응용에 대한 연구가 활발 히 진행되고 있다.

그러나 하이퍼스펙트럴 영상은 많은 밴드수로 인하여 신뢰도가 낮은 트레이닝 데이터를 취득할 경우 영상을 분 류하는데 있어서 오히려 낮은 분류 정확도를 나타낼 수 있다. 따라서 기존의 멀티스펙트럴 영상에 사용되던 일반

2006년 11월 1일 접수, 2006년 12월 4일 채택

^{*} 주저자·서울대학교 공과대학 지구환경시스템공학부 공간정보연구실 박사과정 (choijw11@snu.ac.kr)

^{**} 서울대학교 공과대학 지구환경시스템공학부 공간정보연구실 박사과정 (kko071@snu.ac.kr)

^{***} 서울대학교 공과대학 지구환경시스템공학부 부교수 (yik@snu.ac.kr)

^{****} 서울대학교 공과대학 지구환경시스템공학부 조교수 (kiyun@snu.ac.kr)

적인 분류기법들은 하이퍼스펙트럴 영상 분류에 활용하 는데 어려움을 겪을 수 있다. 따라서 이러한 문제를 해결 하기 위해 최근에는 하이퍼스펙트럴 데이터에 효과적인 새로운 분류 알고리즘에 대한 연구가 활발히 진행되고 있 는데, 그 중 SVM(Support Vector Machine)을 이용한 분 류에 관한 연구가 많은 관심을 받고 있다. SVM은 Vapnik (1995)가 제안한 통계 학습 이론(Statistical learning theory) 에 기반한 알고리즘으로 이분류와 회귀분석 문제를 풀기 위한 방법으로 고안되었다. SVM은 안정적이면서 강력한 일반화 성능을 보인다. 또한, 커널(kernel)을 이용하면 클 래스 간의 비선형 분리가 가능하기 때문에 데이터 마이닝 분야를 비롯한 많은 과학분야에 사용되고 있는 방법이다. 이에 따라 인공위성 영상에 SVM을 이용한 분류기법에 대한 연구가 활발히 진행되고 있다. 먼저, Huang(2002)는 트레이닝 데이터 추출 방법에 따른 SVM 분류기법과 일 반적인 영상분류기법과의 비교를 통해 SVM이 뛰어난 분 류정확도를 나타낸다는 것을 보이고, SVM 분류기 내에서 다양한 변수의 변화를 통해 위성영상의 분류에 적합한 변 수를 도출하였다. 또한, Heng-nian(2004)는 Landsat TM 영상에 SVM 분류기와 fuzzy C-means clustering 알고리 즘을 혼합한 분류방법을 제시하였다. SVM을 하이퍼스펙 트럴 영상에 적용한 연구로는 PAL(2005)이 하이퍼스펙트 럴 영상의 밴드 수 변화에 따른 분류정확도 비교 평가에 관한 연구를 수행하였으며, Mercier(2003)은 하이퍼스펙트 럴 영상의 분광특성을 고려한 분광기반 커널(spectral-based kernel)을 제시하였다.

위의 연구들은 대부분 기존의 SVM을 위성영상에 적용 한 사례이다. SVM 분류에서 하이퍼스펙트럴 영상의 특징 인 높은 분광해상도를 적절하게 이용하지 못하고 위성영 상에의 적용사례에 그치고 있다. 특히 SVM분류기를 설정 할 때 데이터에 맞는 커널을 설정하여야 하는데, 이에 대 한 연구는 부족한 실정이다. Mercier(2003)가 이러한 문제 를 해결하고자 하였지만, 제안한 커널의 경우, 높은 정확 도를 보이고 있지 않다. 따라서, 본 연구에는 하이퍼스펙 트럴 영상에 적합한 분광유사도 기반의 커널을 제시하고 이를 SVM에 적용해 보고자 한다. 이를 위해서 NASA EO-1 Hyperion영상에 실제 제안한 커널과 일반적인 커널 로 SVM분류를 시행하여 보고, 이에 따른 분류정확도를 비교평가해보고자 한다.

2. Support Vector Machine(SVM)

2.1 SVM의 기본 원리

SVM 방법은 이분류를 위해 고안된 방법으로 비선형 패 턴 인식 문제, 함수 회귀문제, 데이터 마이닝, 컴퓨터 비전,

第14卷 第4號 2006年 12月

인공지능, 예측, 의학진단 등의 분야에서 매우 활발하게 연구가 진행되고 있다. 이는 구조적 위험 최소화(structural risk minimization)원리에 기반하고 있어 테스트 단계에서 의 우수한 일반화 성능을 보장하는 학습 알고리즘(training algorithm)이다.

다차원의 공간상에서 학습 데이터(training data)가 $\{x_i, y_i\}$, $i=1,...,r, y_i \in +1,-1$ 와 같이 나타내어질 때, 두 개의 클 래스를 구분하는 초평면(hyperplane)은 여러 개가 있지만, 최적의 초평면(optimal hyperplane)은 그림 1과 같이 하나 만 존재한다.

이러한 최적의 초평면은 각 클래스 데이터 중에서 분리 하는 초평면에 가장 가까운 데이터 사이의 거리를 최대화 할 수 있어야 한다. 최적의 초평면은 그림 1과 같으며 식 (1)에 의해서 정의된다.

$$(\vec{\omega} \cdot \vec{x}) + b = 0 \tag{1}$$

여기서, *x*는 초평면 상의 한 점, *w*는 초평면에서의 법 선이며 *b*는 편향이다.

그림 1과 같이 데이터가 선형분리 가능한 경우에는 두 개의 클래스를 정의하는 초평면은 식 (2)와 같이 정의내릴 수 있으며 이러한 두 개의 초평면 상의 학습 데이터는 support vector라 부른다.

$$y_i(\vec{\omega} \cdot \vec{x_i} + b) - 1 \ge 0 \tag{2}$$

또한, 두 개의 초평면은 두 평면 사이의 마진(margin) 2/(ألَّىالَ)을 최대화하여야 하므로, 두 클래스의 초평면을 구 하기 위해서는 식 (2)를 제약식으로 가지는 다음 목적식의 최적화 문제가 된다.



그림 1. SVM 내에서의 최적의 초평면

$$\min\left\{ \begin{array}{c} \frac{1}{2} \|\vec{\omega}\|^2 \right\} \tag{3}$$

하지만, 대개의 경우 데이터는 선형 분리 가능하지 않기 때문에 위의 제약식을 만족하지 않는다. 이를 해결하기 위 해 초평면으로부터 잘못된 쪽에 위치한 데이터까지의 거 리를 나타내는 여유변수(slack variable)과 패널티 항을 첨 가하여 식 (4)를 제약조건으로 하는 식 (5)의 최적화문제 로 변환할 수 있다.

$$y(\vec{\omega} \cdot \vec{x_i} + b) > 1 - \xi_i \tag{4}$$

$$\min\left[\frac{\|\vec{\omega}\|^2}{2} + C\sum_{i=1}^r \xi_i\right] \tag{5}$$

여기서, ξ_i 는 slack variable, C는 penalty term이다.

2.2 비선형 분리를 위한 SVM

앞선 경우의 SVM은 데이터의 경계가 비선형인 경우에 는 적절한 초평면을 생성할 수 없다. 선형분리 가능한 경 우의 식 (4), 식 (5)를 풀기 위해서는 최적화 과정에서 학 습데이터는 내적의 형태로 나타내어진다. 따라서 비선형 데이터를 특정한 함수에 의해서 입력공간(input space)에 서의 데이터를 고차원의 특징 공간(feature space)으로 사 상 시킨다면 학습 알고리즘 상에서의 내적은 특징 공간에 서의 함수 ϕ 의 내적 $\phi(x_i) \cdot \phi(x_j)$ 으로 표현이 가능하게 된다. 이러한 데이터 공간의 변환을 통해서 비선형 경계를 가지고 있는 데이터의 초평면을 구할 수 있다. 하지만, $\phi(x)$ 를 직접 계산하는 것은 불가능하기 때문에 $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ 를 만족하는 임의의 커널함수(kernel function) K를 사용 하게 된다. 여기서 커널함수는 mercer's condition을 만족 하여야 한다. SVM에서 사용하는 대표적인 커널 함수는 아래의 3가지 함수가 주로 사용된다.

1) polynomial kernel

$$K(x_i, x_j) = (x_i \bullet x_j + 1)^p$$



그림 2. one-against-rest SVM

- 2) RBF(Radial Basis Function) kernel $K(x_i, x_j) = e^{-\gamma |x_i x_j|^2}$
- Hyperbolic tangent kernel K(x_i,x_i) = tanh(kx_i • x_i - δ)

2.3 Multi-class SVM classifier

SVM 분류기는 두 개의 클래스를 분류하는 이분류 문제 에서만 적용이 가능하다. 따라서 다중 클래스에 SVM을 적용하기 위해서는 다른 방법을 사용하여야 한다. 가장 대 표적인 방법으로는 다음의 2가지 방법이 있다.

2.3.1 one-against-rest SVM

이 방법은 먼저, 분류하고자 하는 클래스의 개수만큼의 SVM 분류기를 구축하는 방법이다. i번째 분류기는 i번째 클래스와 나머지 클래스로 할당된 모든 데이터간의 결정 경계를 생성하게 되고, 생성된 결정경계를 이용한 결정함 수에 데이터를 입력하여 가장 큰 함수 값을 가지는 클래 스로 레이블링하는 기법이다.

2.3.2 one-against-one SVM

이 기법은 2개의 클래스로부터의 데이터에 대하여 학습 을 한 총 $\frac{k(k-1)}{2}$ 개의 분류기를 구축하여 각 분류기에 데 이터를 입력하여 테스트하여 가장 빈도가 높은 클래스로 할 당하는 방법이다. 이 방법은 일반적으로 one-against-rest SVM에 비해서 빠른 속도를 보인다고 알려져 있으나, 분류 기가 잘 조정되어 있지 않으며 과적합(overfitting)되는 문제 를 가지고 있다.

2.4 분광 유사도 커널(Spectral Similarity Kernel)

Sweet(2004)는 하이퍼스펙트럴 영상분류에서 두 벡터 간의 거리만을 이용하여 유사도를 측정할 경우에 유사한 특성을 가지고 있는 벡터들을 다른 특성을 가지고 있는 벡터들로 오분류하는 경우가 발생한다는 것을 증명하였 다. SVM에서 비선형분리문제를 해결하기 위해서 적용한



그림 3. one-against-one SVM

韓國地形空間情報學會誌

커널들도 대부분 거리를 기반한 함수를 사용하고 있는데, 이를 하이퍼스펙트럴 영상에 적용할 경우에 위와 같은 문 제가 발생할 수 있다. 따라서 본 논문에서는 RBF 커널을 기본으로 하며, 분광정보의 거리 유사도 뿐만 아니라, 각 유사도를 측정할 수 있는 분광 유사도 커널을 제안한다. 기본 가정은 RBF 커널에 분광각 유사도를 첨가하는 것이 며, 이를 식으로 표현하면 다음과 같다.

$$K(x_i, x_j) = \exp\{-\gamma(||x_i - x_j||^2 + \arccos(\frac{x_i \cdot x_j}{||x_i||||x_j||}))\}$$
(6)

여기서, $\|x_i - x_j\|^2$ 은 거리 유사도척도 이며 $\arccos\left(\frac{x_i \cdot x_j}{\|x_i\| \|x_j\|}\right)$ 는 각 유사도척도를 의미한다. 하지만 2개의 척도는 서로 스케

일이 다르기 때문에 다음의 변형된 식을 통해서 2개 척도의 스 케일을 조정하기로 한다.

$$K(x_i, x_j) = \exp\{-\gamma(||x'_i - x'_j||^2 + S\arccos(x'_i \cdot x'_j))\}$$
(7)

여기서, $x' = \frac{x}{\|x\|}$ 인 단위벡터, *S*는 scale factor이다. 위 의 커널을 통해서 특정 분광정보간의 분광유사도를 SVM 에 적용시킬 수 있다.

3. 알고리즘의 적용 및 결과

3.1 사용 데이터

EO-1 위성은 NASA에서 추진하는 NMP(New Millennium Program)의 첫 번째 지구관측위성으로 2000년 11월 21일에 발사되었다. 사용한 자료는 2002년 4월 3일에 취득한 영상으로 안양과 군포 일대의 250pixels * 500 pixels 영상을 대상지 역으로 사용하여 알고리즘을 평가하기로 하였다.

3.2 실험방법

실험방법은 크게 전처리 과정, training data 추출, SVM 분류기 적용에 따른 평가로 나뉘어 지며 전체적인 흐름도 는 그림 4와 같다.

3.2.1 전처리

정확한 분류를 위해서는 영상에 존재하는 노이즈 자료들 을 제거시켜 주어야 하는데, 이를 위해서 MNF Transform 을 이용하여 노이즈 자료들을 제거하여 주었으며, 이를 기 반으로 수작업을 통하여 기초적인 band selection 작업을 수행하였다.

(1) MNF(Minimum Noise Fraction) 변환MNF 변환은 주성분분석(PCA) 변환과 유사한 변환이



그림 4. 실험의 순서도

다. 그러나, 주성분분석에서 자료들의 산포도를 최대화하 는 주성분으로 변화하는데 반해, MNF변환은 신호대 잡음 비(SNR)를 최대화하는 성분을 선정하는 선형변환이다. MNF변환 후에 생성되는 낮은 순서의 성분들에는 증가된 잡음이 존재하며, 필터링 등의 후처리를 실시하고 이를 MNF 역변환을 실시하면 자료의 노이즈가 감소하는 효과 를 얻을 수 있다.

(2) Band Selection

MNF를 사용한 후에도 영상 전체가 의미없는 정보를 담 고 있는 밴드를 처리를 할 수 없다. 특히, Hyperion 영상에 는 이러한 영상들이 존재하는데, 총 242개의 영상 내에서, 전혀 정보를 담지 않은 영상과, 노이즈로만 채워져 있는 영상으로 구분할 수 있다. 이를 제거하기 위해서는 다양한 알고리즘을 적용할 수 있지만, 본 연구에서는 USGS에서 공식적으로 보정되지 않은 밴드로 명명한 밴드 44개와 수 작업을 통하여 노이즈라고 판단되는 영상 63개를 제거하 였으며, 결과적으로 총 135개의 밴드만을 추출하였다.

(3) Reflectance Transformation

추출된 135개의 밴드 자료는 모두 16bit의 Radiance값 이다. 영상의 Radiance 값은 각 파장에 대하여 지표에서 반사된 에너지의 양이기 때문에 밴드별로 스케일이 모두 다르다. 따라서 이들 값들은 지표의 분광반사 특성 값들로 변화시켜주어야 분류의 성능을 향상시킬 수 있다. 또한, 변환된 분광반사값들은 0~1의 값을 가지기 때문에 SVM 을 적용하기 위해서는 필수적인 사항이라고 할 수 있다. 분광반사로의 변환식은 식 8과 같다.

$$p_P = \frac{\pi \times L_\lambda \times d^2}{ESUN_\lambda \times \cos\theta_s} \tag{8}$$

여기서,

 ρ_P:
 영상의 반사율

 L_λ:
 영상의 분광 복사휘도

 d:
 지구와 태양사이의 거리

 ESUN_λ:
 센서의 태양 방사도

 θ_s:
 태양 천정 각이다.

3.2.2 Training Data 추출

Training Data는 영상에서 영상 내에서 화소들을 직접 추출 하고, class label를 수동으로 입력하여 분류를 위한 Data Set 을 만들었다. 클래스는 그림 5와 같이 forest, shadow-forest, water, city, dark city, field 총 6개의 클래스로 설정하였으며, Data Set을 샘플링 하여 SVM의 학습을 위한 Training Data 1202개와 검증을 위한 Validattion Data 708개를 생성하였다. 정성적인 평가를 위해서는 전체 영상에 대하여 분류기를 적 용하기로 하였다.

3.3 결과

SVM의 경우 polynomial 커널과 RBF 커널, 분광유사도 커널을 사용하였으며 SVM을 적용하는데 사용하는 변수의 경우, penalty term과 slack variable는 공통된 값을 산정하



그림 5. 각 클래스 별 Training Data 추출

였고, 각 커널의 변수는 타 연구(PAL. M, 2005; Heng-nian Q, 2004)를 참고하여 과적합하지 않는 최적의 변수를 사용 하였다.

먼저, 각 커널을 사용하여 Training Data를 이용하여 분 류기를 구성하고, 이를 바탕으로 각 Validation Data의 클 래스와 SVM 분류기를 적용했을 때의 클래스의 일치 여 부를 비교하여 비율인 분류정확도는 표 1과 같다.

표 1에서 확인할 수 있듯이, polynomial 커널보다는 RBF 커널과 분광유사도 커널이 더욱 높은 정확도를 나타내는 것을 확인할 수 있다. 또한, 미세한 양이지만 분광유사도 커널이 RBF 커널보다 정확도가 높은 것도 볼 수 있다. 하 지만, RBF 커널과 본 논문에서 제안한 커널의 정확도 차이 는 통계적으로 유의하다고 볼 수 없다. 이는 Validation Data를 영상 전체에 걸쳐서 획득한 것이 아니라 수작업을 통하여 국부적으로 획득하였기 때문에 다양한 지역에 대 해서 성능을 제대로 평가할 수 없었기 때문으로 판단되어 진다. 이를 보완하기 위하여 Training Data로 학습을 하여 분류기를 구축하고 이를 전체 지역에 대해서 분류를 시행 해 보았다. 결과는 그림 6과 같다.

전체적으로 분광유사도 커널을 사용한 SVM이 영상의 토지피복상태를 가장 잘 표현한다는 것을 확인해 볼 수 있었다. polynomial 커널을 사용하였을 때는 도심지간의 분류가 제대로 이루어지지 않았고, RBF 커널의 경우에는 나대지를 도심지로 오분류하는 지형이 많이 발생하였다. 세부적으로 살펴보면 골프장 주변(그림 7)에서는 기존의 커널을 사용한 SVM이 토지의 상태를 잘 표현하지 못하 고, 골프장 주변을 도심지로 오인하는 화소들이 많이 분포 하는 반면에 본 논문에서 제안한 분광 유사도 커널은 골 프장의 형태를 효과적으로 분류하는 것을 볼 수 있으며, 오분류하는 화소들도 가장 적은 것을 확인하였다. 그림 8 의 도심지 부문에서는 RBF 커널과 분광유사도 커널이 높 은 분광반사값을 가지는 도심지와 높은 반사값을 가지는 도심지를 효과적으로 구분하며 분광유사도 커널은 특히 도심지를 과대분류하는 가능성이 적음을 볼 수 있다. 이를 통해, 분광유사도 커널이 각과 거리척도를 모두 사용하기 때문에, 각 화소들이 분광차를 효과적으로 판별할 수 있다 는 것을 증명할 수 있다.

표 1. SVM의 커널에 따른 분류결과

| 분류정확도 |
|--------|
| 93.78% |
| 96.32% |
| 96.46% |
| |



(a) : Hyperion data



(b) : polynomial kernel



(d) : spectral similarity kernel





그림 8. 커널에 따른 SVM분류기를 사용한 분류결과(도심지 주변)

4. 결 론

본 논문에서는 SVM 분류기를 구성할 때 기존의 커널이 가지는 한계점을 극복하기 위해서 거리 유사도와 각 유사 도, 2개의 척도를 모두 고려하는 분광유사도 커널을 제안 하였다. 제안한 커널을 Hyperion 영상에 적용해 본 결과, 정량적으로는 분광유사도 커널이 가장 뛰어난 성능을 보 임을 확인하였다. 또한, 공간적으로도 본 논문에서 제시한 분광유사도 커널을 이용한 SVM 분류기가 가장 효과적으 로 하이퍼스펙트럴 영상을 분류함을 확인할 수 있었다. 특 히, 도심지 분류에 있어서 도로를 비롯한 낮은 분광 반사 값을 가지는 도심지와 상대적으로 높은 분광 반사값을 가 지는 도심지를 효과적으로 구분해내었다. 골프장과 같은 식물 혼재지역에서도 제안한 커널을 이용하였을 때, 영상 의 공간적 특징을 그대로 가지는 분류결과를 보였다.

본 연구를 통해서 토지피복 분류에 있어서 SVM의 뛰어 난 성능을 검증해 볼 수 있었으며, 분광유사도 커널을 사 용할 때, 기존의 SVM을 더욱 효과적으로 적용할 수 있음 을 보였다. 향후에는 다양한 항목의 분류 및 지형도를 활 용하여 더욱 정밀한 정확도 평가를 수행해야 할 것이며, 분광유사도 커널의 변수 변화에 따른 성능 평가도 수행하 여야 할 것이다.

참고문헌

- Heng-nian Q, Jian-gang Yang, Yi-wen Zhong and Chao Deng, 2004, Multi-class svm based remote sensing image classification and its semi-supervised improvement scheme, *IEEE*.
- Huang, C., DAVIS, L.S. and TOWNSHEND, J.R.G., 2002, An assessment of support vector machines for land cover classification, *International Journal of Remote Sensing*, 23, pp. 725–749.
- Mercier, Gregoir and Lennon Marc, 2003, Support Vector Machines for Hyperspectral Image Classification with Spectral– based kernels, IEEE
- PAL, M. and MATHER, P.M. 2005, Support vector machines for classification in remote sensing, *International Journal of Remote Sensing*, 26, pp. 1007–1011.
- Sweet, James Norman, 2004, The Spectral Similarity Scale and its Application to the Classification of Hyperspectral Remote Sensing Data, *IEEE*.
- Vapnik, V, 1995, The Nature of Statistical Learning Theory, Springer – Verlag, NewYork.