

# 지시벡토르분류방법의 정확도개선에 대한 연구

김 혜 옥

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《나라의 과학기술을 세계적수준에 올려세우자면 발전된 과학기술을 받아들이는것과 함께 새로운 과학기술분야를 개척하고 그 성과를 인민경제에 적극 받아들여야 합니다.》

(《김정일선집》 증보판 제11권 138~139페이지)

최근에 위성화상분류에서는 지시벡토르에 의한 분류방법[1-3]을 많이 리용하고있는데 이 방법은 다른 분류방법들에 비해 분류정확도가 높은것으로 알려져있다. 그러나 현실에서는 더 높은 분류정확도를 요구하는데 이 방법도 역시 표준지자료들이 지형과 위치에 따라 각이한것으로 하여 현실의 요구를 만족시키지 못한다.

논문에서는 모호K평균무리분석법으로 표준지자료를 몇개의 무리로 갈라놓은 다음 지시벡토르에 의한 분류방법으로 분류하여 그것의 정확도를 평가한데 대하여 서술하였다.

## 1. 모호K평균무리분석법에 의한 표준지가르기

각종 지물류형들의 밝음도에는 스펙트르적으로 모호성이 존재하므로 정의한 가격함수를 최소로 하는 모호K평균무리분석법으로 표준지자료들에 대한 무리분석을 진행하여 표준지를 분류한다.

그 절차는 다음과 같다.[1]

① 매 대역의 히스토그램에 기초하여 최빈수를 계산하고 그것들의 조합으로 초기무리중심을 선택하며 모호K평균알고리즘으로 초기모호결수  $u_{li}$ 를 계산한다.

$$E(C) = \sum_{l=1}^L \sum_{i=1}^n u_{li} \|G_i - m_l\|^2$$

여기서  $E(C)$ 는 가격함수,  $G$ 는 표준지들의 측정값으로서  $G = \{G_i | i = \overline{1, n}\}$ ,  $C$ 는 류형들의 모임으로서  $C = \{C_1, \dots, C_L\}$ 이다.

② 무리중심  $m_l$ , 공분산행렬  $v_l$ 을 계산한다.

$$m_l = \sum_{i=1}^n u_{li} G_i / \sum_{i=1}^n u_{li}, v_l = \sum_{i=1}^n u_{li} (G_i - m_l)(G_i - m_l)^T / \sum_{i=1}^n u_{li}, l = \overline{1, L}$$

③ 새 류형의 모호결수  $\bar{u}_{li}$ 를 계산하며  $u = \{u_{l1}, u_{l2}, \dots, u_{ln}\}$ 의 렬들을 표준화한다.

모호결수  $u_{li}$ 를 가격함수식이 최소가 되도록 라그랑주의 미정승수법을 리용한다.

$$u_{li} = (1 / \|G_i - m_l\|^2)^{1/(q-1)} / \sum_{j=1}^L (1 / \|G_i - m_j\|^2)^{1/(q-1)}, j = \overline{1, L}, i = \overline{1, n}$$

표준화식은 다음과 같다.

$$u_{li} = \begin{cases} 1, & i \in C_l \\ 0, & i \notin C_l \end{cases}, \sum_{l=1}^L u_{li} = 1, \quad l = \overline{1, L}, \quad \sum_{i=1}^n u_{li} > 0, \quad i = \overline{1, n}$$

④  $u$ 의 변화가 없을 때까지 ②, ③을 반복한다.

## 2. 무리에 의한 지시벡토르분류방법(SVM법)의 정확도평가

SVM은 표준지자료의 스펙트르공간에서 분포를 규정하지 않고 그에 맞게 류형들의 경계를 구획화하여 분류하며 적은 표준지자료를 가지고도 분류할수 있는 우점을 가지고있다.

여기에서는 선형결심함수  $f(x)=wx+b(w \in \mathbb{R}^N$ 으로서 분리초평면을 결정하는 량,  $b \in \mathbb{R}$ 로서 편위량)를 결정하는것이 기본문제이다.

최량초평면을 얻기 위하여 다음의 제한을 가진 최량문제를 풀어야 한다.

$$\|w\|^2 / 2 + C \sum_{i=1}^k \xi_i \rightarrow \min$$

$$y_i(wx+b) \geq 1 - \xi_i, \quad i = \overline{1, k}$$

식에서  $0 < C < \infty$ 는 정규파라메터이다.

비선형변환으로 핵함수  $K$ 를 얻는다.

$$K(x_i, x_j) \equiv \phi(x_i)\phi(x_j)$$

분류결심규칙은 다음과 같다.

$$f(x) = \text{sign}\left(\sum_{\text{sup port vector}} y_i \alpha_i^0 K(x_i, x) + b^0\right)$$

론문에서는 시지역에 대하여 산림류형을 분류하고 여러 분류방법들과 비교평가하였다.

위성화상자료는 2015년 5월 30일

Landsat 8호 OLI의 117-031, 117-032구역자료이다.

표준지는 삼송나무, 이깔나무, 소나무, 잣나무, 참나무, 넓은잎나무림으로 덮여있다. 우리는 표준지외에 물, 주민지, 밭, 불탄자리, 논, 무림지들에 대하여 위성화상에서 육안판단으로 표준지를 더 선택하였다. 일반적으로 위성자료분류에서 많이 리용하는 방법인 최우도방법과 지시벡토르분류방법 그리고 론문에서 제기한 방법에 의한 분류정확도를 비교한다. 분류정확도는 분류에 리용된 표준지자료에 대하여 분류의 총정확도로 평가한다.(표) 표에서 보는바와 같이 표준지를 그대로 리용하는 경우 SVM법에 의한 분류정확도는 최우도분류에 비

표. 여러 분류방법들에 의한 총정확도(%)

지표	방법		
	최우도 분류방법	SVM법	무리에 의한 SVM법
삼송나무	57.9	87.0	82.9
이깔나무	63.7	99.5	94.7
소나무	79.2	63.0	89.0
잣나무	71.0	5.3	30.8
참나무	82.2	98.0	96.4
기타 넓은 잎나무	20.8	49.8	66.7
물	99.3	100.0	100.0
주민지	100.0	99.6	100.0
밭	100.0	100.0	99.9
불탄자리	99.5	100.0	99.5
무림목지	88.6	80.4	94.9
논	99.4	100.0	100.0
총평균값	80.1	81.9	87.9

해 크게 높아지지 않으며 표준지자료를 무리분류한 경우에 크게 높아진다는것을 알수 있다.

## 맺는 말

론문에서는 류형들의 표준지자료에 대하여 모호K평균무리분석법으로 무리를 구성하고 지지벡토르분류방법으로 분류할 때 분류정확도가 크게 높아진다는것을 확정하였다.

## 참고 문헌

- [1] J. C. Morton; Image Analysis, Classification and Change Detection in Remote Sensing, Taylor & Francis, 233~244, 2007.
- [2] W. Pakorn et al.; Photogrammetric Engineering & Remote Sensing 1, 74, 2, 239, 2008.
- [3] A. R. S. Marcal et al.; International Journal of Remote Sensing 1, 26, 7, 1347, 2005.

주체106(2017)년 8월 5일 원고접수

## Accuracy Improvement of SVM Method

*Kim Hye Ok*

We analysed that dividing sampling data of classes had into some kinds of classes an influence on the accuracy improment by the SVM.

On the basis of this analysis, we suggested the accuracy improvement method of SVM by the fuzzy K-means clustering.

Key words: SVM, fuzzy K-means clustering, optimal hyperplane