空間データのカテゴリ比率の半教師あり推定
(Semi-supervised Contextual Unmixing of Geospatial Data)

尾崎智彦[1]　　西井 龍映[2]　　澤村 陽子[1]

[1] 九州大学 大学院数理学府，[2] 九州大学 大学院数理学研究院

1 Introduction

Suppose that multispectral geostatistical data, i.e. feature vectors are observed at respective pixels, are given. Contextual image classification is a problem for allocation of pixels to one of land-cover categories by learning the feature vectors as well as adjacency relationship of the pixels. Image classification is an important and fundamental issue in remotely-sensed data analysis for environmental studies, e.g., making land-cover maps for a wide variety of purposes. It is known that contextual classification methods based on Markov random fields (MRF) improve non-contextual classifiers successfully. Even though training data are unavailable, contextual clustering methods still work well.

Now, consider the situation such that a low-spatial resolution multispectral image is given, and we are required to estimate fractions of categories in each mixed cell (**mixel**). This issue is called **unmixing**. [**?**] consider contextual unmixing methods without training data for inferring the categories. Our aim here is to extend their contextual unmixing method with small training data.

2 Assumptions for contextual unmixing

Suppose that there are K land-cover categories $C_1, ..., C_K$ in the image, and feature vector $\boldsymbol{X}$ from C_k follows a d-dimensional normal distribution $N_d(\boldsymbol{\mu}_k, \Sigma_k)$ with mean vector $\boldsymbol{\mu}_k$ and variance-covariance matrix Σ_k for $k = 1, ..., K$. Suppose that a training data set $\{(\boldsymbol{x}_s^*, y_s^*) | s = 1, ..., m\}$ is available, where $y_s^* \in \{1, ..., K\}$ (**pure pixels**). In addition to the training set, a test data set consisting $n (>> m)$ **mixels** without their labels is also given. Let $\boldsymbol{x}_i$ be a feature vector at test mixel $i = 1, ..., n$, and $\boldsymbol{f}_i = (f_{i1}, ..., f_{iK})$ be a vector of fractions where K categories cover mixel i. Note that $f_{ik} \geq 0$ and $\sum_{k=1}^K f_{ik} = 1$. Our contextual unmixing method is based on the following two assumptions.

Assumption 1: (Conditional independence and normal mixtures)
Assume that all feature vectors given all fraction vectors are spatially independent, and each conditional distribution follows a Gaussian mixture. This implies $p(\boldsymbol{x}_1, ..., \boldsymbol{x}_n | \boldsymbol{f}_1, ..., \boldsymbol{f}_n) = \prod_{i=1}^n p(\boldsymbol{x}_i | \boldsymbol{f}_i) = \prod_{i=1}^n \left\{ \sum_{k=1}^K f_{ik} \phi(\boldsymbol{x}; \boldsymbol{\mu}_k, \Sigma_k) \right\}$, where $\phi(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma)$ denotes a probability density function of $N_d(\boldsymbol{\mu}, \Sigma)$.

Assumption 2: (MRF for category fractions)
Consider mixel i and its four neighbors, see Fig. 1. We assume that fraction vectors follow MRF, i.e., the conditional probability density of the fraction vector at i given all fractions except itself is expressed by

$$p\big(\boldsymbol{f}_i | \boldsymbol{f}_j, j \neq i\big) \propto \exp\left(-\beta \|\boldsymbol{f}_i - \overline{\boldsymbol{f}}_i\|^2\right) \times \frac{\Gamma(\alpha_1 + \cdots + \alpha_d)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_d)} \prod_{\ell=1}^d f_{i\ell}^{\alpha_\ell - 1}$$

where $\alpha_j \geqq 1$ are constants and $\overline{\boldsymbol{f}}_i \equiv (\boldsymbol{f}_{iN} + \boldsymbol{f}_{iS} + \boldsymbol{f}_{iE} + \boldsymbol{f}_{iW})/4$.

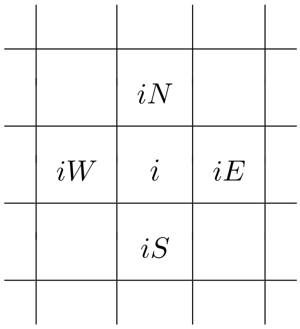

Fig. 1: Pixel i and its neighbors

3 Semi-supervised contextual unmixing

Our main concern is to maximize the posterior density $p(\boldsymbol{f}_1, ..., \boldsymbol{f}_n | \boldsymbol{x}_1, ..., \boldsymbol{x}_n)$. This problem is, however, hard because the joint distribution of MRF cannot be expressed in the closed form. So, we maximize pseudo-conditional likelihoods, and the fractions are estimated. Now, we propose the following target function to be maximized:

$$\sum_{s=1}^{m} \log \phi(\boldsymbol{x}_s^*; \widehat{\boldsymbol{\mu}}_{y_s^*}, \widehat{\Sigma}_{y_s^*}) + \lambda \sum_{i=1}^{n} \log p(\boldsymbol{x}_i | \boldsymbol{f}_1, ..., \boldsymbol{f}_n)\,,$$

where $0 < \lambda < 1$ is a tuning parameter. The function is maximized by an EM algorithm-like procedure. The proposed method is examined through artificial and benchmark data sets for supervised classification, and it shows a satisfactory performance.

References

R. Nishii et al. (2008). Contextual unmixing of geospatial data based on Gaussian mixture models and Markov random fields. *IEEE Proc. IGARSS 2008.*