Application of spatial classification rules for remotely sensed images

Giedrius Stabingis, Lijana Stabingienė

Department of Mathematics and Statistics, Klaipėda University

H. Manto 84, LT-92294 Klaipeda

E-mail: gstabingis@gmail.com, lijana.stabingiene@gmail.com

Abstract. In this paper the remote sensed image classification example using spacial classification rule with distance (SCRD) is examined. This supervised classification method was first presented in paper [11]. This method is improved version of earlier method PBDF [4, 10, 9], during the classification it incorporates more spatial information. The advantage of this method is its ability to classify data which is corrupted by Gaussian random field and it is typical to remotely sensed images classified in this letter which are corrupted by clouds. Classification accuracy is compared with earlier method and with other commonly used supervised classification methods.

Keywords: image classification, spatial classification rules, supervised classification.

Introduction

The classification of remotely sensed images is widely used in different situations. Lots of methods performs very good in clear situations, but often the remotely sensed images or some parts of them are covered with clouds. In this case these common methods perform errors, because they don't have the information behind the clouds. The methods authors are working with helps in this exact situation, the clouds are modeled according to the training sample information and spatial relation between observation to be classified and the training sample, and this modeled information is used during classification for better decision making.

The earlier method [4, 10] was tested in the real situation for classification of the real remotely sensed image covered with clouds in article [9]. Later the more advanced method was created [11] and now in this letter the performance of this method is tested in the same situation. In the paper [9] the remotely sensed image was used from Landast7 satellite. Now, when Landsat8 satellite images are available, the same situation was tested on Landsat8 satellite image which is more corrupted by clouds and the image is larger when the first one.

In order to show the complicity of such situation some other common methods are used in the same situation for classification. For the commonly used supervised classification methods the rassclass package from the R system is used. Numerical and visual analysis of discriminant function in the case of isotropic exponential spatial correlation for the nearest neighbor neighborhood system using eight nearest neighbors $NN(8)$ is done. All calculations are done in R system.

1   Method declaration

The incorporation of the spatial information into image classification is done by some authors (see e.g., [1, 2]). The statistical supervised classification method (PBDF) [4] was extended by incorporating more influence from the spatial dependency into classification problem and achieving higher accuracy in the paper [11] and this method is named SCRD.

The plugin Bayes discriminant function (PBDF) to the classification problem is

$$W_t(Z_0; \hat{\Psi}) = \left(Z_0 - \frac{(\hat{\mu}_{1t}^0 + \hat{\mu}_{2t}^0)}{2} \right) \frac{(\hat{\mu}_{1t}^0 - \hat{\mu}_{2t}^0)}{\hat{\sigma}_{0t}^2} + \gamma(y), \tag{1}$$

where $\hat{\mu}_{lt}^0 = \hat{\mu}_l + \alpha_0'(z_n - X_y\hat{\mu})$, and $\hat{\sigma}_{0t}^2 = \hat{\sigma}^2 R_{on}$, $\gamma(y) = \ln(\pi_1(y)/\pi_2(y))$.

The PBDF classification rule based on the posterior distribution of $Y(s_0)$ specified by

$$\pi_1(y) = \frac{1}{1 + \exp\left\{ \rho\left(1 - \frac{2n_1}{n}\right) \right\}}, \quad n_1 = 0, 1, \dots, n, \tag{2}$$

when $I_0 = \{i\colon s_i \in N_0 = NN(8),\ i = 1, \dots, n\}$ and where ρ is non negative constant called a clustering parameter, and n_1 is the number of locations from N_0 with label equal 1. Here $NN(8)$ is the nearest neighbor scheme with eight nearest neighbors.

The new classification rule SCRD is based on the following posterior distribution

$$\pi_1(y) = \frac{\sum_{i \in I_0} \frac{\delta(y_i=1)}{d(s_i, s_0)}}{\sum_{i \in I_0} \frac{1}{d(s_i, s_0)}}, \tag{3}$$

where $\delta(\cdot)$ is the 0–1 indicator function and $d(\cdot, \cdot)$ denotes the Euclidean distance function between locations. For the case of two classes $\pi_2 = 1 - \pi_1$.

2   Methods for comparison

In order to verify the reliability of the method and its advantage in this concrete situation, when data is corrupted by spatial noise, the comparison with other commonly used methods is done. These methods are: Support Vector Machines (SVM), Neural Networks (NNet), Random Forests (RF) and Multinomial Logistic Regression (Logit). All these methods belong to the group of supervised classification methods and the classification with these methods is done on the same data and on the same training sample as for the methods proposed in this letter.

The advantage of SVMs classifiers is their capabilities to learn from small number of samples [5]. The class (label) of a new sample is determined by a linear combination of the kernel functions evaluated on a certain subset of the examples the support vectors and the input. The coefficients of the combination are obtained as a solution to a convex optimization problem occurring at the learning stage [8].

Neural networks rely on the iterative derivation of weights which effectively define hyper-planes and hyper-regions in the pattern feature space [7]. Although artificial neural network methods are frequently found to give a higher total classification accuracy when compared to other methods, they do not always perform universally well [6].
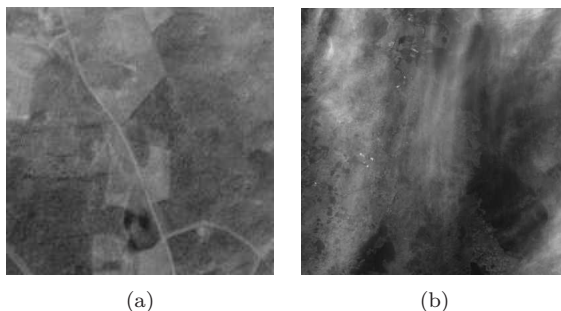
(a) (b)

Fig. 1. Images used for classification.

Random Forests method grows many classification trees and then every tree gives the class label for the observation to be classified. Then the class is assigned according to the most classes given by all trees in the forest [3].

3   Description of experiment

In this letter the experiment with two different remotely sensed images is presented. One image was earlier used in the paper [9]. Here the same image is classified with the new method. This image is obtained by Landsat7 satellite and is from the territory of Lithuania. This image is 200×200 pixels and is shown in Fig. 1(a). Another image is obtained by the satellite Landsat8 and is also in the territory of Lithuania. The second image is 4 times bigger: 400×400 pixels. This image is shown in Fig. 1(b). The density of clouds is higher in the second picture so the accuracy of classification must be lower for this image.

The original image is naturally corrupted with clouds and such noise is modeled by Gaussian random field (GRF) with zero mean and exponential spatial correlation function given by $r(h) = \exp\{-|h|^2/\alpha\}$. Here α is a spatial correlation range parameter which must be estimated. This parameter is evaluated with function variofit from the geoR package in the environment of the R system.

The remotely sensed image used for classification is naturally corrupted and then the exact correlation range parameter is unknown. This parameter is estimated according to the training sample points (pixels) using the geoR package of the statistical program R.

At the same time both classes training sample points are used in order to identify the correlation range parameter. In order to minimize the influence of different classes to the accuracy of the modeling, the means of corresponding classes are subtracted from the training sample points feature values. This way transformed points with their coordinates are then used to calculate empirical semivariogram. It is done with geoR packages command variog. The parametric model is then fitted to the empirical semivariogram points using variofit command, which uses least squares method for fitting. After several trials, several different types of models were fitted according to the shape of the empirical semivariogram, it was determined that for this concrete situation best model was of exponential type. The exponential model was best fitted for the both images. The value of the correlation range parameter $\alpha = 13$ for the first image and $\alpha = 68$ for the second image were estimated. These numbers were used for classification using SCRD and PBDF methods.
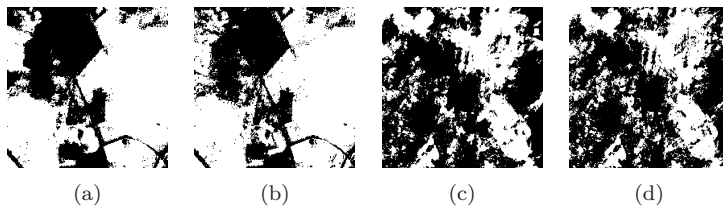
Fig. 2. Classification with Bayes discriminant function results. (a), (c) with SCDR; (b), (d) with PBDF.
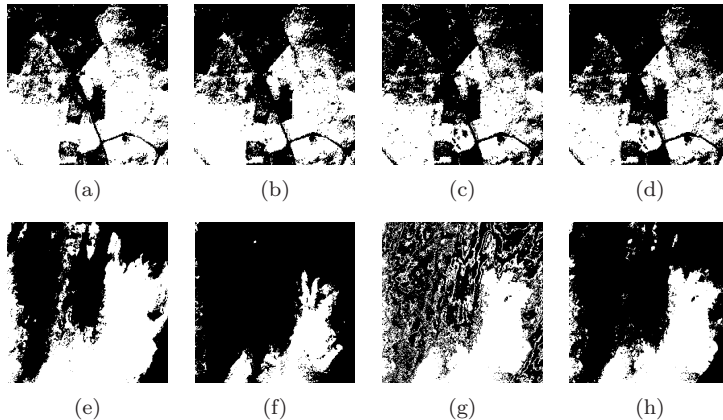

Fig. 3. Classification results with common supervised classification methods. (a), (e) with Logit. (b), (f) with NNet. (c), (g) with RF. (d), (h) with SVM.

Table 1. Classification accuracy.

	SCRD	PBDF	Logit	NNet	RF	SVM
Image Fig. 1(a)	0.908	0.882	0.757	0.757	0.695	0.696
Image Fig. 1(b)	0.869	0.842	0.596	0.678	0.578	0.652

4   Results

Visual classification results are shown in Fig. 2 and in Fig. 3.

Numerical evaluations of overall accuracy for all methods used in this experiment are provided in Table 1.

5   Conclusions

The SCRD method performed better than PBDF, so the incorporation of more spatial information into classification rule helps to get better results.

The second image is more corrupted by clouds than the first one and all methods performed worse.

If the clouds corrupt image just slightly as in the situation with the first image, other common methods still performs quite well, but if corruption becomes harder, these methods become useless.

SCRD and PBDF methods performs quite well in both situations, even when image is corrupted quite hard.

References

[1] P.M. Atkinson. Spatially weighted supervised classification for remote sensing. *Int. J. Appl. Earth Observ. Geoinf.*, **5**:277–291, 2004.

[2] P.M. Atkinson. A geostatistically weighted k-NN classifier for remotely sensed imagery. *Geograph. Anal.*, **42**:204–225, 2010.

[3] L. Breiman. Random forests. *Mach. Learn.*, **45**(1):5–32, 2001.

[4] K. Dučinskas, L. Stabingienė and G. Stabingis. Image classification based on bayes discriminant functions. In *Proceedings of the 1th International Conference "Spatial Statistics 2011-Mapping Global Change"*, 7, pp. 218–223, 2011.

[5] A.K. Jain, P.W. Duin and J. Mao. Statistical pattern recognition: a review. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22**(1):4–37, 2000.

[6] I. Kanellopoulos, G.G. Wilkinson and J. Megier. Integration of neural network and statistical image classification for land cover mapping. *Geosc. Rem. Sens. Symp.*, 1993.

[7] R.P. Lippmann. An introduction to computing with neural nets. *IEEE ASSP Mag.*, **40**:4–22, 1987.

[8] W.W.Y. Ng, A. Dorado, D.S. Yeung, W. Pedrycz and E. Izquierdo. Image classification with the use of radial basis function neural networks and the minimization of the localized generalization error. *Patt. Recogn.*, **40**:19–32, 2007.

[9] L. Stabingienė. Classification of the real remotely sensed image covered with clouds. *Liet. matem. rink. Proc. LMS, Ser. A*, **53**:117–112, 2012.

[10] L. Stabingienė. *Image analysis using Bayes discriminant functions.* Doctoral dissertation. Vilnius University, Vilnius, 2012.

[11] G. Stabingis, K. Dučinskas and L. Stabingienė. Comparison of spatial classification rules with different conditional distributions of class label. *Nonl. Anal. Mod. Contr.*, **19**(1):109–117, 2014.

REZIUMĖ

Erdvinių klasifikavimo taisyklių taikymas nuotolinio stebėjimo vaizdams
G. Stabingis, L. Stabingienė

Šiame straipsnyje pateikiamas erdvinės klasifikavimo taisyklės (SCRD) taikymas nuotolinio stebėjimo vaizdų klasifikavimui. Šis klasifikavimo su mokymu metodas buvo pristatytas straipsnyje [11]. Šis metodas yra patobulinta ankstesnio metodo [4, 10, 9] PBDF versija, klasifikavimo metu dar labiau atsižvelgiant į erdvinės informacijos įtaką. Šio metodo privalumas yra galimybė klasifikuoti duomenis, kurie yra sugadinti Gauso atsitiktiniu lauku, o tai yra būdinga šiame darbe klasifikuojamiems vaizdams iš palydovo, kuriuose tiriama teritorija yra uždengta debesimis. Klasifikavimo tikslumas yra lyginamas su ankstesniu metodu ir kitais plačiai taikomais klasifikavimo su mokymu metodais.

Raktiniai žodžiai: vaizdų klasifikaimas, erdvinės klasifikavimo taisyklės, klasifikavimas su mokymu.