Neighborhood scheme selection for classification with SCRD method

Giedrius Stabingis[1,2], Lijana Stabingienė[2]

[1]*Institute of Mathematics and Informatics, Vilnius University*
  Akademijos 4, LT-08663 Vilnius

[2]*Department of Computer Science and Statistics, Klaipeda University*
  H. Manto 84, LT-92294 Klaipeda

  E-mail: giedrius.stabingis@mii.vu.lt, lijana.stabingiene@gmail.com

Abstract. When using Spatial Correlation Rule with Distance (SCRD) the selection of the neighborhood scheme influences classification accuracy. Spatial dependency in different situations remains at various distances, so, according to this, in applications it is important to choose a suitable neighborhood scheme. In the earlier papers of the authors, the nearest neighbor scheme was used. In this paper, several different neighborhood schemes are examined by large experiment.

Keywords: spatial classification, supervised classification, neighborhood, neighborhood schemes.

Introduction

Spatial information is important in image classification [2]. In the papers (see e.g., [1]) the incorporation of geostatistical information of features into plug-in versions of classifiers is based on the marginal distribution of the observation to be classified. The statistical supervised classification method was extended by incorporating more influence from the spatial dependency into classification problem in the paper [3] and this method is named SCRD. The SCRD method is used in situations when classification data is with spatially correlated information (noise). This is common in images which are covered with clouds or smoke in pollution data and in other situations.

SCRD method showed its better accuracy comparing with other common methods in [3] in artificial experiment and in the paper [4] in real life situation. Analyzing more deeply the results of the previous experiments some important cases were noticed. During the classification some parts were misclassified because there were no training samples corresponding appropriate class. These parts were surrounded by the training sample points of other class. In this letter this situation is examined by proposing to use different neighborhood schemes which always includes at least one training sample element from each class.

The declaration of SCRD method and descriptions of neighborhood schemes are presented in the first section. In the second section the experiment is described. In the third section the numerical and visual results of the experiment are presented.

1   Method description

The method used in this letter is a spatial classification rule based on the Plug-in Bayes Discriminant Dunction (PBDF) with posterior distribution of class label depending on distances among unclassified locations and training sample locations is called SCRD [3]. As it is common in image analysis the features are modeled by stationary Gaussian Random Field (GRF) $\{Z(s)\colon s \in D \subset \mathbb{R}^2\}$, and class labels are modeled by discrete Markov Random Field (MRF). Here s is a state of the pixel.

$Z(s) = \mu_l + \varepsilon(s)$ is the marginal model of observation $Z(s)$ in class Ω_l, with the mean μ_l and with the error term $\varepsilon(s)$ which is generated by zero-mean stationary GRF $\{\varepsilon(s)\colon s \in D\}$ with covariance function defined by model $cov\{\varepsilon(s), \varepsilon(u)\} = \sigma^2 r(s - u)$ for all $s, u \in D$, where σ^2 is variance as a scale parameter. In this letter the exponential covariance function is used $C(h) = \sigma^2 \exp\{-|h|/\alpha\}$. $r(s - u) = r(h) = \exp\{-|h|/\alpha\}$ is the spatial correlation function, where α is the correlation range parameter which shows how far the correlation remains and h is the Euclidean distance between s and u locations.

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

$$\pi_1(y) = \left(\sum_{i \in I_0} \frac{\delta(y_i = 1)}{d(s_i, s_0)} \right) \Big/ \left(\sum_{i \in I_0} \frac{1}{d(s_i, s_0)} \right), \tag{2}$$

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 $\pi_2 = 1 - \pi_1$.

$I_0 = \{i\colon s_i \in N_0, \ i = 1, \ldots, n\}$ and n_1 is the number of locations from N_0 with label equal 1. Here N_0 is a set of s_0 states neighboring pixel selected by a neighborhood scheme.

We test eight different neighborhood schemes which are using four different methods for selecting neighbors from training sample. The differences between neighborhood schemes are presented in Fig. 1.

$\texttt{NN}(n)$ method selects n nearest neighbors form the training sample. According to this method, various amounts of different class training sample elements can be selected. There can be a situation when all selected neighbors are from the same class so the classified pixel is also assigned to this class.

$\texttt{NN}_C(n, m)$ method selects n nearest neighbors form the training sample as the $\texttt{NN}(n)$ method. If n_i (selected neighboring points from the i-th class) is smaller then m, then additional $m - n_i$ neighbors are selected from the training sample from the i-th class and it is done for all classes. In this case there is always at least m elements from all classes and this means that the information from all classes is used for classification for every pixel to be classified.

$\texttt{NN}_{R1}(rad, n)$ method selects neighbors form the training sample where Euclidean distance between training sample pixels and pixel to be classified is smaller or equal
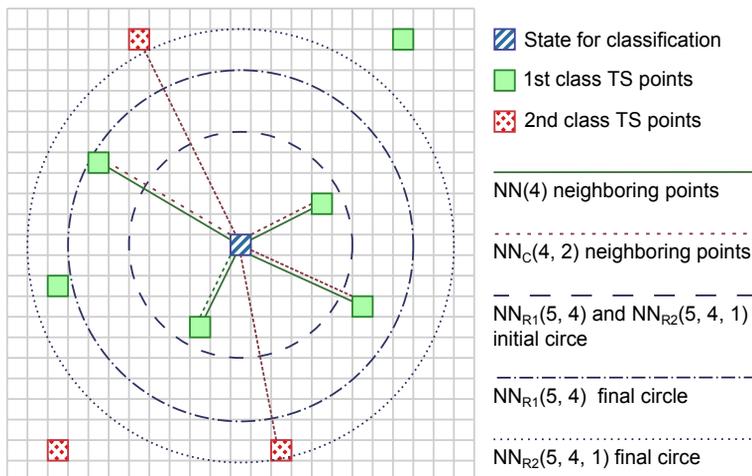
**Fig. 1.** Examples of neighborhood schemes used in this letter.

to the radius rad. In this case the situation can occur when no trainings sample is selected. So if the number of selected neighbors from the training sample is smaller then n, the radius rad is increased until at least n neighbors are selected. Using this method the same situation as with $\texttt{NN}(n)$ method can occur, when all selected neighbors are from the same class, so another modification is needed.

$\texttt{NN}_{R2}(rad, n, m)$ method selects neighbors form the training sample in the same way as $\texttt{NN}_{R1}(rad, n)$ method. When at least n neighbors are selected, the radius rad is increased until m training sample elements are selected for all classes.

2 Description of experiment

In order to determine the influence of neighborhood schemes a large experiment was performed. 100 different initial black and white images of 200×200 px size were made. White color of the images corresponds to the first class and the black color – to the second class. From the initial images training sample points were taken randomly selecting about 0.8% of the points for the training sample. Then two dimensional GRF images were generated using exponential covariance function with different correlation range parameter α which varied from 5 to 60 by 5. So for each initial image 12 different GRF images were generated. Then these GRF images were multiplied by 3 and added with the initial images, then the received images were normed. So 1200 different images were generated for the classification. The preparation of the experiment scheme is presented in Fig. 2.

The images for the classification were classified with SCRD method using neighborhood schemes: $\texttt{NN}(4)$, $\texttt{NN}(8)$, $\texttt{NN}_C(2,2)$, $\texttt{NN}_C(4,2)$, $\texttt{NN}_{R1}(10,4)$, $\texttt{NN}_{R1}(30,4)$, $\texttt{NN}_{R2}(10, 4,1)$, $\texttt{NN}_{R2}(20,4,1)$. The classification accuracy (ac) was calculated $ac = N_C/N$, where N_C is the number of correctly classified pixels and $N = 40\,000$ is the number of pixels per image. All calculation were done with statistical software R [5].
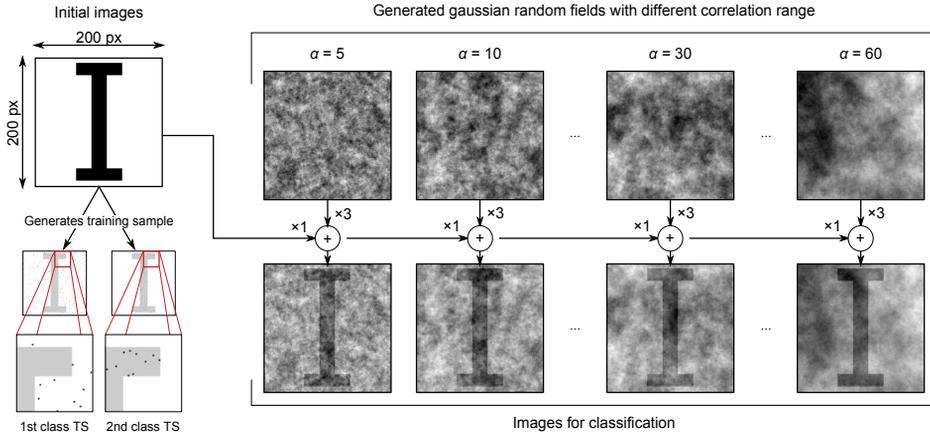
Fig. 2. Preparation of the experiment. 100 different initial images we processed according to this scheme and 1200 different images for the classification were generated.

Table 1. Average classification accuracy for different neighborhood schemes using SCRD method.

Method	$\mathrm{NN}(4)$	$\mathrm{NN}(8)$	$\mathrm{NN}_C(2,2)$	$\mathrm{NN}_C(4,2)$
ac	98.36	98.96	98.36	98.60
Method	$\mathrm{NN}_{R1}(10,4)$	$\mathrm{NN}_{R1}(30,4)$	$\mathrm{NN}_{R2}(10,4,1)$	$\mathrm{NN}_{R2}(30,4,1)$
ac	98.80	98.88	98.61	98.92

3 Results

After the classification the results were received. Table 1 shows the average accuracy of the classification using SCRD method for all neighborhood schemes used in this experiment. According to the results presented in Table 1, all neighborhood schemes were almost equally good.

The influence of the correlation range to the classification accuracy is presented in Fig. 3. As seen in the graph all neighborhood schemes performs better when correlation range parameter α increases and it almost stops increasing at $\alpha = 35$. From Table 1 and from Fig. 3 can be seen that schemes $\mathrm{NN}(8)$, $\mathrm{NN}_{R1}(30,4)$ and $\mathrm{NN}_{R2}(30,4,1)$ performs best but the accuracy difference is very small.

Visual results of most interesting and important images are presented in Fig. 4. These visual results show that best schemes, according to average classification accuracy, in some cases does not deal with harder situations when where are no training samples in specific places. The image of Fig. 4 representing number "87" is the best example. Only the methods which uses at lest one element form each training sample could correctly classify the bottom of the number "8".

Conclusions

Neighborhood schemes which require to take some training samples from every class $\mathrm{NN}_C(n,m)$ and $\mathrm{NN}_{R2}(rad,n,m)$ deals better with harder situations, when in some
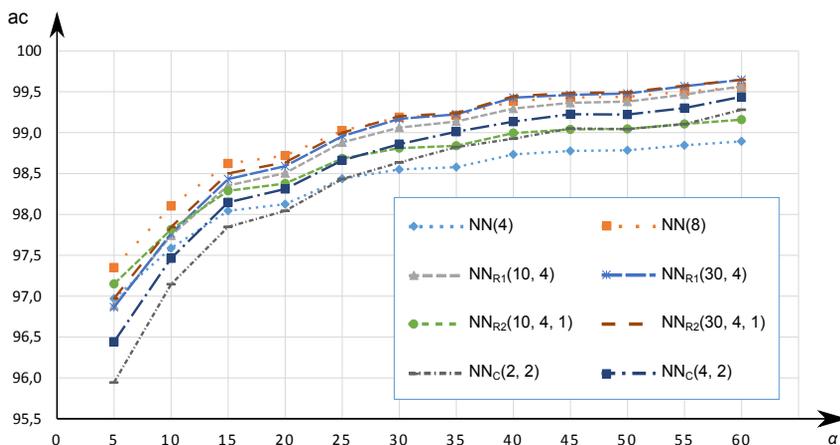
Fig. 3. Results of the classification accuracy (ac) using SCRD classification method with different neighborhood schemes. Results are presented according to different correlation range parameter α values.
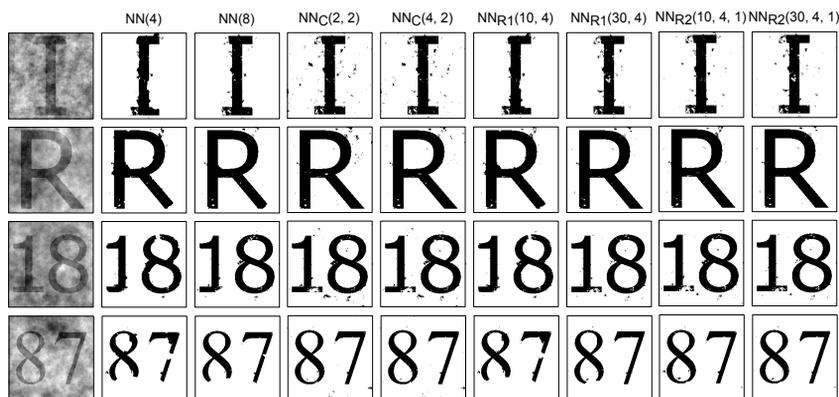

Fig. 4. Visual results of the classification. Most interesting images are presented.

important place training samples are missing. It may became more important when the smaller training sample is given. Such schemes can help in situations when some class areas are very thin and the training samples can not be gain from these thin areas.

These different neighborhood schemes can be tested in real situation, especially classifying roads or rivers from remotely sensed images.

$NN_C(n, m)$ and $NN_{R2}(rad, n, m)$ neighborhood schemes during the classification gets more noisy results then other methods, but this noise can be removed using common morphological image processing methods.

References

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

[2] 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.

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

[4] G. Stabingis and L. Stabingienė. Application of spatial classification rules for remotely sensed images. *Liet. matem. rink. Proc. LMD, Ser. B*, **55**:63–67, 2014.

[5] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. Available from Internet: http://www.R-project.org/.

REZIUMĖ

Kaimynystės schemos parinkimas klasifikuojant SCRD metodu
G. Stabingis ir L. Stabingienė

Naudojant erdvinio klasifikavimo taisyklę su atstumu (SCRD) rezultatų tikslumą įtakoja kaimynystės schemos parinkimas. Erdvinė priklausomybė skirtingose situacijose išlieka skirtingais atstumais, todėl taikymuose, atsižvelgiant į tai, svarbu yra pasirinkti tinkamą kaimynystės schemą. Ankstesniuose autorių darbuose būdavo naudojama artimiausių kaimynų schema. Šiame straipsnyje tiriamas keleto skirtingų kaimynystės schemų veikimas atliekant platų eksperimentą.

Raktiniai žodžiai: erdvinis klasifikavimas, klasifikavimas su mokymu, kaimynystė, kaimynystės schemos.