An improved $k$-NN method based on multiple-point statistics for classification of high-spatial resolution imagery

Y Tang, L Jing, H Li, Q Liu and H Ding

Key Laboratory of Digital Earth Sciences, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, 100094, China
E-mail: tangyunwei2008@163.com

Abstract. In this paper, the potential of multiple-point statistics (MPS) for object-based classification is explored using a modified $k$-nearest neighbour ($k$-NN) classification method (MP$k$-NN). The method first utilises a training image derived from a classified map to characterise the spatial correlation between multiple points of land cover classes, overcoming the limitations of two-point geostatistical methods, and then the spatial information in the form of multiple-point probability is incorporated into the $k$-NN classifier. The remotely sensed image of an IKONOS subscene of the Beijing urban area was selected to evaluate the method. The image was object-based classified using the MP$k$-NN method and several alternatives, including the traditional $k$-NN, the geostatistically weighted $k$-NN, the Bayesian method, the decision tree classifier (DTC), and the support vector machine classifier (SVM). It was demonstrated that the MP$k$-NN approach can achieve greater classification accuracy relative to the alternatives, which are 82.05% and 89.12% based on pixel and object testing data, respectively. Thus, the proposed method is appropriate for object-based classification.

1. Introduction

It is generally agreed that object-based image analysis (OBIA) has several advantages compared to pixel-based analysis for processing fine spatial resolution remotely sensed data [1]. Some advanced algorithms for object-based classification have been developed. For example, Chen et al. [2] introduced a modified object-based method integrating multiple characteristics based on fuzzy logic. Salehi et al. [3] developed a hierarchical rule-based, object-based classification framework coupled with height data for classification in complex urban environments. Object properties and spatial context are important for OBIA classification in remote sensing. Commonly, a spatial weighting is incorporated into the classifier to increase the accuracy of the classification result. It has been suggested that such contextual classifiers can reduce noise and achieve greater accuracy than non-contextual classifiers [4]. Geostatistical techniques have been applied to pixel-based classifiers as a contextual classification method [5]. However, all the spatial weighting schemes investigated were based on two-point statistics (i.e., the relationship between the central pixel and its neighbours).

Recently, advances in multiple-point statistics (MPS) have shown that two-point statistics are limited. Instead of two-point-based functions such as the variogram, MPS borrows structures from training images, from which higher-order local patterns of the target field can be captured. Thus, spatial structures can be characterised more completely using MPS [6]. The original implementation of the MPS approach builds on the paradigm of the traditional indicator simulation. Recently, a few studies applied MPS to the classification of remotely sensing data. For example, Boucher [7] realised super-resolution mapping with MPS. Ge and Bai [8] extracted linear objects from remotely sensed...
imagery using MPS. Tang et al. [9] proposed a post-classification method using MPS and compared it with contextual classification based on the Markov random field model. Nevertheless, MPS for the classification of fine spatial resolution imagery remains under-explored.

The goals of this paper are: (i) development of a new MPS spatial weighting for remotely sensed contextual classification, and (ii) evaluation of the applicability of the new method for handling objects (c.f., the previous pixel-based implementation). The multiple-point weighted k-NN method is introduced in Section 2, along with a general description of MPS, and a set of standard k-NN classification methods. Our experiment in the study area is presented in Sections 3. Section 4 provides further discussion on the proposed method, followed by some concluding remarks.

2. Methods

In the traditional k-NN method, the classifier allocates pixels to the neighbours to which it is closest in feature space. An inverse distance weighting (IDW) function can be incorporated into the k-NN classifier to give more weight to information from a neighbour close to an unclassified observation than from a more distant neighbour. IDW can be expressed as:

$$\omega_{uk} = \frac{1}{d_{uk}^p}$$  \hspace{1cm} (1)

where $d_{uk}$ measures the distance between the current pixel $u$ and its neighbouring training pixel $k$ in feature space, $\omega_{uk}$ is the weight based on an inverse distance, and the exponent $p$ is an integer that determines the magnitude of the weight. The term wk-NN is used to refer to the IDW-based k-NN method.

In a geostatistically weighted k-NN classifier (gk-NN), the probability that a pixel $u$ belongs to class $m$ can be evaluated as follows [5]:

$$p_{gk-NN}[c(u) = m] = \frac{\sum_{k=1}^{K} S_g \times p_{m,m} (h_{uk}) \times \omega_{uk} + (1-S_g) \times \omega_{uk}}{\sum_{m=1}^{M} \sum_{k=1}^{K} S_g \times p_{m,m} (h_{uk}) \times \omega_{uk} + (1-S_g) \times \omega_{uk}}$$  \hspace{1cm} (2)

where the subscript $uk$ of $h$ indicates the lag between pixel $u$ and its neighbour $k$. $p_{m,m} (h_{uk})$ is the fitted model of the spatial covariance, which also refers to the class-conditional probability. The term $m'$ is a class index for $m' = 1, \ldots, M$ classes, and $m$ is the class of interest. $S_g$ is a proportional weight between 0 and 1.

In the proposed MPS-based k-NN approach, instead of training samples, the conditional probability is derived from the training image to provide spatial information. For an unknown location $u$, $K$ nearest neighbour training pixels $u_k$ ($k = 1, \ldots, K$) can be found. Thus, the data template at location $u$ can be defined as $T(u) = \{h_{u1}, \ldots, h_{uk}\}$, where $h_{uk}$ is the lag between $u_k$ and $u$. So the template centred at $u$ consists of the same separation lag (for both distance and direction) and classes with the neighbouring $K$ pixels. This template is used to scan the training image and derive the multiple-point probability for pixel $u$ by counting the replicates of the data events $dev(u)$, where $dev(u) = \{c(u_1), \ldots, c(u_K)\}$. The training image can be obtained from an initial classification map. The probability of pixel $u$ with class $m$ equals the proportion of the number of $dev(u)$ that possesses class $m$ at the central node to the total number of $dev(u)$. For another pixel, a different template is applied to estimate another probability from the training image. The multiple-point probability that a pixel $u$ belongs to class $m$ is, thus, expressed as:

$$p_{mp}[c(u) = m] = \frac{\sum_{u=1}^{N} I[c(u) = m | dev(u) = 1]}{\sum_{u=1}^{N} I[dev(u) = 1]}$$  \hspace{1cm} (3)
Note that $dev(u) = 1$ means that the data event $dev(u)$ is found in the training image, indicating that all $K$ pixels at location $u_k$ should match exactly the corresponding classes (i.e., $c(u_k) = m_k$ ($k = 1, \ldots, K$)). Thus, the multiple-point based $k$-NN classifier (MP$k$-NN in brief) across $L$ probability levels can be written as:

$$p_{\text{MP}\text{-NN}}[c(u) = m] = S_{\text{MP}} \times p_{\text{MP}}[c(u) = m] + (1 - S_{\text{MP}}) \times p_{\text{gl}\text{-NN}}[c(u) = m]$$

$$= S_{\text{MP}} \times \frac{1}{L} \sum_{l=1}^{L} \sum_{u=1}^{N} \frac{\sum_{u=1}^{N} I[c(u) = m \, \mid \, dev^l(u) = 1]}{\sum_{u=1}^{N} I[dev^l(u) = 1]}$$

$$+ (1 - S_{\text{MP}}) \times \frac{\sum_{k=1}^{K} \sum_{m=1}^{M} S_{g} \times p_{m,m}(h_{ak}) \times \omega_{ak} + (1 - S_{g}) \times \omega_{ak}}{\sum_{m=1}^{M} \sum_{k=1}^{K} S_{g} \times p_{m,m}(h_{ak}) \times \omega_{ak} + (1 - S_{g}) \times \omega_{ak}}$$

(4)

Similar to $S_g$, $S_{\text{MP}}$ is a multiple-point statistical weight given to the classifier, ranging from 0 to 1.

To estimate the multiple-point probability in the MP$k$-NN method, a training image is defined first. It can be obtained either for the same area that needs to be classified using the MP$k$-NN method or represent a different place. However, the training image is required to have a similar class spatial distribution to that of the study area and, thus, to provide prior knowledge on the character of spatial information. Then the data template is defined, which consists of $K$ nearest neighbour pixels of the current pixel. Therefore, one pixel in the image corresponds to only one data template, and the data templates are different at each location. The multi-grid concept is applied to the data template. Instead of expanding the data template, the rescaled template is constructed by condensing the original one. Thus, the data template is formulated as: $T^l(u) = \left\{ h_{i}/2^{l-1}, \ldots, h_{n}/2^{l-1} \right\}$.

To apply the MP$k$-NN method, the difference between the pixel- and object-based methods is the measure of spatial distance when modeling the spatial correlation. When the pixel-based method is applied, the distance is measured directly between two points. For the object-based method, the point located at the center of each segment is extracted first. The central locations are recorded, so that the distance between any two segments can be inferred. It should be noted that, here, the $k$-NNs are segments, and the distances from the current segment to the $k$-NNs are measures of the relations between the centers of the segments. However, the scanning process is pixel-based, from top to bottom and left to right of the whole training image. The replicates of the data events were recorded according to the class type of the central node. This information was then converted to a conditional probability for each class and incorporated into the $gk$-NN classifier, as shown in equation (4).

3. Case study: IKONOS image of Beijing

3.1. Study area and data processing

The MP$k$-NN method proposed in this paper was tested on a 256 by 256 pixel large subset (39°57‘55“–39°58‘28“N, 116°24‘5“–116°24‘49“E) of an IKONOS image acquired over the Beijing urban area in May of 2000 (figure 1a). The dataset consists of four multispectral bands with a spatial resolution of 4 m. The typical urban area (including that in the present study) can be classified generally into four classes: buildings, vegetation, road/bare land and shadow. The focus is mainly on distinguishing the building class from bare land/road. Because these two classes, which are key for urban planning and
resource management, have similar spectral responses, but different spatial distributions, their separation is challenging using traditional classifiers.

Reference data were provided to test the classification result. The reference image was produced using eCognition software. The standard \(k\)-NN classifier was applied to the object features of brightness, maximum difference, mean values, and the ratio of length to width of the four spectral bands. Considering that some bias may be introduced in the classification process, the classification result was then modified manually with reference to a corresponding image in Google Earth, which was obtained in January of 2001. The final reference image is shown in figure 1b, which has the same segmentation number as the segmentation image to be classified.

The training data were used with two objectives in mind. One was to train the classifier, and the other was to estimate the class-conditional probability for geostatistical modelling. It is common to use objects instead of pixels for object-classification, since the object contains both spectral and shape information. A total of 35 training objects were selected manually from the 1348 possible segments, which are displayed in figure 1c. These training segments were used to train the \(k\)-NN classifier. Although the size of the training set (35 segments relative to the total number of 1348) was appropriate for classification, geostatistical modelling requires more sample data. Furthermore, the training segments are not suitable for geostatistical modelling as the model of spatial covariance is a function of distance. Each segment accounts for an area, so the distance measured between segments varies. Therefore, instead of segments, sample points were used for geostatistical modelling.

A stratified sampling scheme was applied to the 35 training segments, ensuring there was at least one point within each segment. The number of points sampled from each segment was based on the size of the segment: the larger the segment, the more sample points that were selected. A total of 50 sample points were selected for each class, and 200 points were selected in total for geostatistical modelling, as displayed in figure 1d.

![Figure 1](image-url)

**Figure 1.** Images relating to the study area in Beijing: (a) IKONOS subset (true colour composite, image credit: GeoEye/DigitalGlobe), (b) reference classification image, (c) training segments selected for classification, and (d) spatial distribution of training points for geostatistical modelling (circles, triangles, vertical crosses, and diagonal crosses represent buildings, vegetation, road/bare land, and shadow, respectively).

### 3.2. Classification

The traditional \(k\)-NN classifier was first applied to the 35 training samples, based on the mean value of the four spectral bands of each segment and with \(K = 5\). For the \(gk\)-NN method, class-conditional probability plots were first estimated from the 200 sample points in figure 1d and then fitted with covariance-type models. Anisotropy was not considered. \(gk\)-NN using equation (2) was performed, based on the same training segments as for \(k\)-NN. Conditional probabilities were estimated from the \(k\)-NN segments (\(K = 5\)) for each class. The proportional weight for the geostatistical information \(S_g\) was taken as 0.8, after experimentation. For comparison with other classifiers, three more classification methods were applied as benchmarks: the Bayesian, decision tree classifier (DTC) and support vector
machine (SVM) methods. Finally, the MPk-NN method was applied. Since a pre-classified image of the same area can guarantee a similar spatial distribution of land covers, the training image was derived from an initial classification result. Here, the land cover map derived from wk-NN was used as the training image, which itself was not used as a benchmark method. The data template for each pixel consisted of $k$-NN nodes ($K = 5$). The multi-grid level $L$ was taken as 3. The classification results obtained using the above six methods are displayed in figure 2. Comparing to the reference map in figure 1b, land cover of vegetation and shadow in the classified maps have similar patterns using the three $k$-NN classification methods. Some shadows were misclassified as buildings using the Bayesian method, and the DTC method confused some trees with shadows in the upper left of the image. Most buildings in the SVM result were not classified correctly. For the three $k$-NN methods, more buildings are shown in the upper right part of the image for the $g_k$-NN and MP$k$-NN results than for the $k$-NN result. The buildings in the reference map are of regular shape and more connected geographically. The $g_k$-NN and MP$k$-NN results are closer in appearance to the reference than those for the $k$-NN result. Roads and bare land are easily confused with buildings. Thus, the decrease in the prevalence of this class can be attributed to the fact that the previous misclassified road and bare land features (in the $k$-NN classified map) were corrected as buildings (in the $g_k$-NN and MP$k$-NN classified maps).

![Figure 2. Classification of land cover in Beijing using six methods: (a) Bayesian, (b) DTC, (c) SVM, (d) $k$-NN, (e) $g_k$-NN, and (f) MP$k$-NN (see text for detail).](image)

As seen in the reference map, for the main road near the bottom of the image, there are only a few trees along the road, without buildings or shadow within the road feature. For the same area, shadows and buildings appear in the main road in the Bayesian, $k$-NN, and $g_k$-NN results. Here, the SVM and MP$k$-NN methods show closer results to the reference data than the other methods. However, for the two squares above the main road, the $k$-NN and DTC classifications were more accurate. This is
probably because bare land arranged as a square has similar features to buildings both spectrally and spatially, so the spatial weights involved in the enhanced classification methods are ineffective. In contrast, in the upper right part of the image, the land cover type of buildings was increased in coverage by the \( g_k \)-NN and MP\( k \)-NN methods, and the narrow roads and small areas of bare land that were misclassified by the \( k \)-NN were corrected in the \( g_k \)-NN and MP\( k \)-NN methods.

Table 1. Classification accuracies of the Bayesian, DTC, SVM, \( k \)-NN, \( g_k \)-NN and MP\( k \)-NN methods in Beijing (AA and OA represent the average accuracy and overall accuracy, respectively).

| Method | Class | Classified number | Correct number | AA (%) | OA (%) |
|--------|-------|-------------------|----------------|--------|--------|
|        |       | Pixel | Object | Pixel | Object |       | Pixel | Object |
| Bayes  | 1     | 18790 | 577   | 9165  | 345   | 54.88 | 66.83  | 70.44 | 67.06 |
|        | 2     | 22823 | 381   | 20768 | 334   | 93.55 | 90.09  |       |       |
|        | 3     | 19962 | 327   | 12306 | 163   | 61.35 | 48.28  |       |       |
|        | 4     | 3961  | 63    | 3925  | 62    | 71.99 | 67.34  |       |       |
| DTC    | 1     | 14361 | 328   | 10128 | 247   | 68.96 | 64.10  | 72.36 | 64.02 |
|        | 2     | 16628 | 258   | 14825 | 221   | 78.88 | 73.44  |       |       |
|        | 3     | 23270 | 544   | 15115 | 254   | 69.98 | 59.74  |       |       |
|        | 4     | 11277 | 218   | 7354  | 141   | 74.65 | 73.57  |       |       |
| SVM    | 1     | 4561  | 274   | 1923  | 202   | 27.48 | 58.49  | 71.25 | 69.44 |
|        | 2     | 21145 | 339   | 19985 | 314   | 93.50 | 89.80  |       |       |
|        | 3     | 31275 | 565   | 16577 | 261   | 67.63 | 60.49  |       |       |
|        | 4     | 8555  | 170   | 8207  | 159   | 94.89 | 93.26  |       |       |
| \( k \)-NN | 1 | 17780 | 396 | 12039 | 290 | 73.91 | 67.67 |
|        | 2     | 19647 | 315   | 19259 | 301   | 93.57 | 89.47  | 81.69 | 72.85 |
|        | 3     | 18923 | 452   | 13898 | 229   | 71.20 | 58.14  |       |       |
|        | 4     | 9186  | 185   | 8340  | 162   | 93.08 | 91.15  |       |       |
| \( g_k \)-NN | 1 | 21524 | 549 | 13462 | 408 | 76.06 | 80.84 |
|        | 2     | 19955 | 311   | 19362 | 301   | 93.31 | 89.62  | 80.48 | 79.75 |
|        | 3     | 15305 | 311   | 11753 | 208   | 67.56 | 63.24  |       |       |
|        | 4     | 8752  | 174   | 8165  | 158   | 93.33 | 91.60  |       |       |
| MP\( k \)-NN | 1 | 17350 | 433 | 13539 | 359 | 84.07 | 79.89 |
|        | 2     | 21177 | 352   | 20488 | 323   | 95.78 | 90.62  | 89.12 | 82.05 |
|        | 3     | 18636 | 399   | 16101 | 264   | 83.15 | 70.91  |       |       |
|        | 4     | 8373  | 164   | 8276  | 160   | 96.74 | 95.56  |       |       |

Table 1 summarises the accuracy results using the six classification methods. The accuracies were estimated based on both pixel-based and object-based testing data. The average accuracy is the average of the user’s accuracy (commission errors) and producer’s accuracy (omission errors). As can be seen, the accuracies of the vegetation class are rather high (all around 90%) except for the DTC method. The SVM and three \( k \)-NN methods resulted in high accuracies for shadows. For the classes of buildings and road and bare land, MP\( k \)-NN has the greatest average accuracies, whereas the SVM and Bayesian methods offered the lowest average accuracies for buildings and road and bare land, respectively. It can be seen that the object-based accuracy does not always accord with the pixel-based accuracy. This is because sometimes one classifier misclassifies some very large segments whereas another classifier misclassifies more but smaller segments. Nevertheless, the MP\( k \)-NN method
produced the greatest overall accuracies relative to the other methods based on both pixel and object testing data. An analysis of variance (ANOVA) for the classification accuracy was performed using the $F$-test. All the classification results were first compared to the reference data to produce a set of indicator values. The value was set to one if the classification code equals the reference class, otherwise zero. This process was applied both on pixels and objects. Then the indicator values were compared with the MP$k$-NN result in turn. As shown in table 2, the $F$-ratios are quite large for the object-based results. The MP$k$-NN result shows significant increases in accuracy with respect to all the other results at the 90% confidence interval using pixel-based assessment. Using object-based assessment, the increases are smaller, but the increase in accuracy for the MP$k$-NN is still significant relative to all the other methods except for the g$k$-NN method.

Table 2. ANOVA for the classification methods with respect to the MP$k$-NN method using $F$-test (Y = yes, N = no).

| Method | Level | $F$-ratio | Significant at 90% level |
|--------|-------|-----------|--------------------------|
| Bayes  | Pixel | 7490.16   | Y                        |
| DTC    | Pixel | 6196.44   | Y                        |
| SVM    | Pixel | 6934.19   | Y                        |
| k-NN   | Pixel | 1466.53   | Y                        |
| g$k$-NN| Pixel | 1925.14   | Y                        |
| Bayes  | Object| 82.15     | Y                        |
| DTC    | Object| 115.91    | Y                        |
| SVM    | Object| 59.59     | Y                        |
| k-NN   | Object| 33.03     | Y                        |
| g$k$-NN| Object| 2.31      | N                        |

4. Discussion and conclusion
Since the classification methods applied are all object-based, it is necessary to explore the applicability of the g$k$-NN and MP$k$-NN methods for object-based classification. For the spatial measures in geographical space in the g$k$-NN method, converting segments to points may lead to some problems. The first problem relates to the sampling density for geostatistical modelling. A stratified sampling scheme was used to estimate the class-conditional probability to ensure that the points have a random spatial distribution, while not too many points were chosen within one segment. Instead of the correlation within one segment, the g$k$-NN method characterises the correlation between segments. Another problem is that modelling the spatial correlation of areas (segments) with representative points may cause the modifiable areal unit problem (MAUP) [10], which is a very common problem of scale when working with spatial correlation. A solution to this problem is to use the area-to-point approach for geostatistical modelling. However, if the study area is very large, and a great number of objects are produced by image segmentation, then geometric features (size, shape, orientation) can be ignored. Therefore, although the g$k$-NN method generally results in a more accurate classification than the k-NN method, it is not ideally suited for object-based classification. The geostatistical modelling process, therefore, requires further exploration.

The research presented here explored the potential of MPS for spatially weighting a remote sensing classifier generally and, in the present case, for object-based classification. A multiple-point statistical k-NN classification method was tested on remotely sensed image of an urban area in Beijing. The multiple-point classification method can account for spatial correlation at multiple points simultaneously, in contrast to common spatial weighting schemes, which are limited to two-point
statistics. The results demonstrate that greater classification accuracy can be achieved using the proposed classification approach compared to the Bayesian, DTC, SVM, traditional K-NN, and gK-NN methods. Future research should be directed towards: (i) analysis of the choice and nature of training imagery; (ii) the method of conversion from objects to points in support of location- and distance-based spatial weighting; (iii) development of a self-adaptive method for parameter estimation and (iv) comparison the MPS-based weighting with other spatial weightings such as conditional random fields (CRF) or Markov random fields (MRF).

References
[1] Blaschke T and Strobl J 2001 What’s wrong with pixels? Some recent developments interfacing remote sensing and GIS GIS – Z. Geoinf. Sys. 14 12-17
[2] Chen Z, Wang G and Liu J 2012 A modified object-oriented classification algorithm and its application in high-resolution remote-sensing imagery Int. J. Remote Sens. 33 3048-3062
[3] Salehi B, Zhang Y, Zhong M and Dey V 2012 Object-based classification of urban areas using VHR imagery and height points ancillary data Remote Sens. 4 2256-2276
[4] Pasolli E, Melgani F, Tuia D, Pacifici F and Emery W J 2014 SVM active learning approach for image classification using spatial information IEEE Trans. Geosc. Remote Sens. 52 2217-2233
[5] Atkinson P M and Naser D K 2010 A geostatistically weighted k-NN classifier for remotely sensed imagery Geogr. Anal. 42 204-225
[6] Guarino F and Srivastava R M 1993 Multivariate geostatistics: beyond bivariate moments Geostatistics Tróia ’92 ed A Soares (Dordrecht: Kluwer Academic Publications) pp 133-144
[7] Boucher A 2009 Sub-pixel mapping of coarse satellite remote sensing images with stochastic simulations from training images Math. Geosci. 41 265-290
[8] Ge Y and Bai H 2011 Multiple-point simulation-based method for extraction of objects with spatial structure from remotely sensed imagery Int. J. Remote Sens. 32 2311-2335
[9] Tang Y, Atkinson P M, Wardrop N A and Zhang J 2013 Multiple-point geostatistical simulation for post-processing a remotely sensed land cover classification Spat. Stat. 5 69-84
[10] Cressie N A 1996 Change of support and the modifiable areal unit problem Geogr. Syst. 3 159-180

Acknowledgements
The publication of this paper was supported by the 100 Talents Program of the Chinese Academy of Science (Y34005101A), the National Natural Science Foundation of China (41501489), and the National Science and Technology Support Program (2013DFG21640).