A study of supervised classification accuracy in fuzzy topological methods

Wenzhong Shi, Kimfung Liu, Hua Zhang

Research Centre for Advanced Spatial Information Technology, Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong
School of Environment Science and Spatial Informatics, China University of Mining & Technology, Xuzhou, China

1. Introduction

Land cover classification accuracy is one of the major research interests in the field of remote sensing (Zhang et al., 2008; Tang et al., 2005). Land cover refers to such ground surface conditions as, forest, grassland, industrial zones, residential zones. Many classification methods have been devised and tested by researchers in an attempt to improve classification accuracy (Shih and Schowengerdt, 1983; Lee and Philpot, 1991; Jakomulska and Stawiecka, 2002; Bruzzone and Carlin, 2006; Berberoglu et al., 2007; Gamba et al., 2007; Tseng et al., 2008), but further overall classification accuracy improvement is still necessary, for fulfilling the requirements of the users who need very high quality land cover classification, such as for a county or national level land resource survey. Thus, in this study, an effective fuzzy topology eigen-value multiple classifier system (MCS) method is provided to improve MCS accuracy. The method is tested by comparing the classification of satellite images for a test area in China with measurements already gained by conventional surveying methods.

Here a brief review of the existing classification methods is given. Three commonly used classifiers are:

(a) the method of maximum likelihood classification (MLC) (Yang, 1993) with posterior probability

\[ P(k|x) = \frac{P(k)P(x|k)}{\sum_{i=1}^{m} P(i)P(x|i)} \]  

where \( P(k) \) is the prior probability of class \( k \), \( P(x|k) \) is the conditional probability to observe \( x \) from class \( k \) (probability density function).

In the case of normal distributions, the likelihood function, \( P(x|k) \), can be expressed as

\[ L_k(x) = \frac{1}{(2\pi)^{n/2}|\Sigma_k|^{1/2}} \exp \left( -\frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) \right), \]

© 2010 Published by Elsevier B.V.
(b) The minimum distance (MIND) (Bow, 1992) classifier with the posterior probability

\[ P_k(x) = \frac{\exp \left\{ -(1/2)(x - \mu_k)^T \sigma^{-1}_k (x - \mu_k) \right\}}{\sum_{j=1}^{m} \exp \left\{ -(1/2)(x - \mu_j)^T \sigma^{-1}_j (x - \mu_j) \right\}}; \]  

and

(c) the Mahalanobis distance classifier (MAH) (Deer et al., 1996) with the posterior probability

\[ P_k(x) = \frac{\exp \left\{ -(1/2)(x - \mu_k)^T \Sigma^{-1}_k (x - \mu_k) \right\}}{\sum_{j=1}^{m} \exp \left\{ -(1/2)(x - \mu_j)^T \Sigma^{-1}_j (x - \mu_j) \right\}}. \]  

The above three classifiers are the most commonly used in the multiple classifier system (MCS). The major objective in combining sets of classifiers is to achieve an accuracy higher than that of the best individual classifier in the set.

Fuzzy topology is a major branch of Mathematics and, many fuzzy sets theories and fuzzy topological theories have been developed (Zadeh, 1965; Chang, 1968; Wong, 1974; Wu and Zheng, 1991; Liu and Luo, 1997). Methods for dealing with fuzzy classification problems have been proposed (Andrew et al., 2001; Bruzzone and Carlin, 2006; Gamba et al., 2007; Joao et al., 2007; Tseng et al., 2008). The creation and subsequent provision of fuzzy topological space is a useful enabling contribution to image classification (Tang et al., 2005). Shi et al. used thresholding to create a fuzzy topological space (Shi and Liu, 2007) and this viewpoint of fuzzy topological space is applied in image segmentation to enhance the classification accuracy. The study presented in this paper introduces fuzzy topology theory to classification method design, thereby treating the boundary and interior of the object, are combined by using the property of spatial connectivity in fuzzy topology.

The rest of the paper is organized as follows. In Section 2, fuzzy topology, induced by dual threshold values, is reviewed for the purpose of understanding the structure of this induced fuzzy topology. Fuzzy topological classification methods used for land classification are reviewed in Section 3. The logical flow of the several classification methods is given in Section 4 and a land cover classification experiment is described in Section 5. Section 6 presents a new method to identify the optimum threshold value and also the accuracy assessment and kappa assessment of the experimental results. Finally, the fundamental structure of this newly introduced method is discussed.

2. Fuzzy topology induced by dual threshold values

An \(I\)-fuzzy subset on \(X\) is a mapping (called the membership function of \(A\)) \(\mu_A : X \rightarrow I\), i.e. the family of all the \([0, 1]\)-fuzzy or \(I\)-fuzzy subsets on \(X\) is just \(I^X\) consisting of all the mappings form \(X\) to \(I\). \(I^X\) here is called an \(I\)-fuzzy space, \(X\) is called the carrier domain of each \(I\)-fuzzy subset of it, and \(I\) is called the value domain of each \(I\)-fuzzy subset of \(X\).

Fuzzy topology is an extension of the ordinary topology. According to Chang (1968), let \(X\) be a non-empty ordinary set, and \(I\) be an \(I\)-lattice. \(\delta\) is called an \(I\)-fuzzy topology on \(X\) if \((X, \delta)\) is called an \(I\)-fuzzy topological space (\(I\)-fts), if \(\delta\) satisfies the following conditions: (i) \(0, 1 \in \delta\); (ii) if \(A, B \in \delta\), then \(A \wedge B \in \delta\); (iii) let \(\{ A_i : i \in J \} \subset \delta\), where \(J\) is an index set, then \(\bigvee_{i \in J} A_i \in \delta\).

Interior, closure and boundary of a fuzzy set are fundamental elementary in fuzzy topological space, where the interior of \(A\) is defined as the join of all the open subset contained in \(A\), denoted by \(A^o\). The closure of \(A\) is defined as the meet of all the closed subset containing \(A\), denoted by \(\bar{A}\). The boundary of \(A\) is defined as \(\partial A = \bar{A} \wedge A^c\).

Moreover, every interior or closure operator can essentially define a fuzzy topology (Liu and Luo, 1997) separately. Based on this understanding, a fuzzy topology can be defined by the interior and closure operators which are, in turn, defined by a suitable level cut. For a fuzzy set (see Fig. 1) in \(X\), the interior operator and closure operator are defined as

\[ A_\alpha(x) = \begin{cases} A(x) & \text{if } A(x) > \alpha \\ 0 & \text{if } A(x) \leq \alpha \end{cases} \]

(see Fig. 2(a)) and

\[ A^{1-\alpha}(x) = A(x) \begin{cases} 1 & \text{if } A(x) \geq 1 - \alpha \\ A(x) & \text{if } A(x) < 1 - \alpha \end{cases} \]

(see Fig. 2(b)) respectively. Finally, the collection of those two sets of interior and closure induced an \(I\)-fuzzy topology \((X, \tau_\alpha, \tau^{1-\alpha})\).
in X, where \( \tau_\alpha = \{ A_\alpha : A \in \mathbb{P} \} \) are the open sets and \( 1^{1-\alpha} = \{ A^{1-\alpha} : A \in \mathbb{P} \} \) is the closed sets. The elements in \( \tau_\alpha \) and \( 1^{1-\alpha} \) satisfy the relations \( (A_\alpha)^\beta = (A^\beta)^{1-\alpha} \), for all of fuzzy set A, i.e., the complement of the elements in the \( \tau_{\alpha_0} \) closed set. Such fuzzy topology is known as dual I-fuzzy topology which is determined by a pair of dual threshold values. The details are given in Liu and Shi (2006).

To measure the magnitude and the direction of the relationship between two or more classes’ distributions, the intercorrelation coefficient of class \( c_i \) and class \( c_j \) is defined as,

\[
r_{ij} = \frac{\sum_{k}(x_k - \mu_{x_i})(y_k - \mu_{y_j})}{\sqrt{\sum_{k}(x_k - \mu_{x_i})^2}(\sum_{k}(y_k - \mu_{y_j})^2)}.
\]

where, the predefined class \( c_i \) with training set \( \{x_1, x_2, \ldots, x_n\} \) and mean is equal to \( \mu_{x_i} \), and the predefined class \( c_j \) with training set \( \{y_1, y_2, \ldots, y_m\} \) and mean is equal to \( \mu_{y_j} \). This coefficient varies between zero and one and the value of this coefficient is closer to one if two distributions are highly correlated. However, the value of the coefficient is closer to zero if two distributions have a lower intercorrelation. Moreover, those intercorrelation coefficients, as stated above, induce a fuzzy topological space. Indeed, suppose there are \( m \) predefined classes, \( c_1, c_2, \ldots, c_m \) and let \( A_1, A_2, \ldots, A_m \) be the corresponding membership functions, for the overall predefined \( m \) classes, the threshold value is then defined as \( \alpha = \max_{i,j} r_{ij} \).

The value \( \alpha = \max_{i,j} r_{ij} \) consequently defines a fuzzy topological space (Liu and Shi, 2006) and each class is then split into three parts: an interior, a boundary and an exterior (see Fig. 3). The interior of class \( c_i \) is denoted by \( (A_\alpha)_i \) and the boundary of \( c_i \) is denoted by \( \partial A_{\alpha_i} = (A_{\alpha_i})^{1-\alpha} \cap (A_{\alpha_i})^\alpha \). The interior part of the class represents the precise part of this class; while the boundary is the imprecise part. This boundary region is reclassified by using connection theory proposed in Liu and Shi (2006). In other words, for each boundary pixel \( x_0 \) searches its 8-connected pixels. If there exist the highest pixel number of a class within these 8-connected pixels, the boundary pixel \( x_0 \) is classified as a null class (see Fig. 4(c)). If the pixel number of classes \( A \) is equal to the pixel number of classes \( C \), then the boundary pixel \( x_0 \) can be either classified as class \( A \) or class \( C \) (see Fig. 4(b)). Repeat this step until no more pixels in the boundary are to be reclassified. By repeating these steps, all the pixels in the boundary are finally reclassified.

Shi et al. (2010) introduced the concept of dual threshold fuzzy topological space to facilitate the decomposition of objects into three parts: an interior, a boundary and an exterior. In the study presented in this paper, an attempt is made to use dual threshold fuzzy topological space to decompose classes into interior, boundary and exterior divisions. The concept of connection in fuzzy topology is then applied to combine the interior and the boundary (Shi and Liu, 2007). The following is the structure of this induced fuzzy topology (Liu and Shi, 2006) for the classification (Fig. 5).

3. Fuzzy topological classification method

A major issue of the traditional classification methods is classification accuracy. In many cases, the classification accuracy can be improved by combining several classifiers with weightings – the multiple classifier systems, where one of the key techniques is to determine an effective combination method that makes use of the benefits of each of the classifiers and avoids the weaknesses of the classifiers. Fumera and Roli (2005) presented a theoretical and experimental analysis of linear combiners for multiple classifier systems. Briem et al. (2002) applied bagging algorithms, boosting algorithms, and consensus-theoretic classifiers to combine multi-source remote sensing and geographic data to improve the accuracy of land cover classification. To further improve the accuracy of the multiple classifications, the concept of fuzzy theory is introduced into the classification scheme. By applying the concept of fuzzy topological space, we intend to solve the problem of mixed pixels to a certain extent, where a pixel may contain several classes with a certain percentage for each of the composed classes. Within a pixel, the percentage of each of the component classes is quantified by the corresponding membership value.

Fuzzy topology is an extension of ordinary topology (Chang, 1968) and provides a natural way of describing the relationship between classes in an image. For \( m \) predefined classes, \( c_1, c_2, \ldots, c_m \) membership functions \( A_{c_1}, A_{c_2}, \ldots, A_{c_m} \) are used to describe the percentages of these classes. For a fuzzy set \( A \), the interior is defined as the junction of all the open subsets contained in \( A \) and denoted by \( A^o \). The closure of \( A \), where all closed subsets meet containing \( A \), is denoted by \( \overline{A} \). The boundary of \( A \) is defined as \( \partial A = \overline{A} \wedge \overline{A} \). According to Liu and Luo (1997), the family of all the fuzzy topologies on \( X \) is one to one corresponding with the family of all respective interior and closure operators. The fuzzy topological space, \( \{ (X_x, \tau_x, 1^{1-\alpha}) : \alpha \in (0, 1) \} \), obtained by Liu and Shi (2006) provides a platform for the decomposition of classes into interior, boundary and exterior and carries the properties of fuzzy topology. That is, a class \( c \) with the membership function \( A_c \) has the interior of \( (A_c)_1 \) and the boundary \( \partial A_c = (A_c)^{1-\alpha} \cap (A_c)^{\alpha} \). In such a case, each class in the image can be regarded as a fuzzy set in a fuzzy topological space. If there are \( m \) predefined classes, then each class

![Fig. 3. The threshold value of four classes.](image)

![Fig. 4. The method of classifying the boundary pixels.](image)
has an individual interior and boundary. The pixels that belong to the interior of a certain class are classified as that particular class definitely, while the pixels that belong to the boundary of that class has to be reclassified. The classification rule for the boundary pixels is based on the connection theory.

Here a brief review on the connection theory is given as follows. Fuzzy topology can be used to study and describe the structure of a neighbourhood and the leveling of spaces. Thus, if the interior and boundary are assumed to have a certain relationship, they can then be reconstructed for certain applications (e.g. for improving classification accuracy) using this relationship. Connectivity is actually one of the features of this relationship. The usual definition of a fuzzy subset, $A$, connection, in fuzzy topology, is that $A$ can not be separated by two non-zero open or closed fuzzy sets, called open connected and closed connected, respectively. In accordance with the characteristics of fuzzy topology, this kind of connection also contains two structural types, neighbourhood and leveling (Liu and Luo, 1997). The connectivity of a spatial object depends on the neighbourhood structure of the object itself, rather than the leveling structure. Thus, the ordinary definition of fuzzy topology connection is not appropriate for spatial object connectivity, and it is necessary to give a new connection definition. Hence, in this study, the concept of spatial object connection is newly defined to indicate whether a spatial object is connected to another spatial object, in the sense of background space, only. This definition is more fully associated with the concept of neighbourhood, rather than the concept of leveling. Therefore, the concept of connection for spatial objects has to be defined on background topology only, making it necessary to define a so called, supported connection. This supported connection depends on the specific background topological space of the fuzzy topological space. Thus, this concept of supported connection can be applied to handle spatial objects with connection properties. For example, the concept of the supported connection is used to classify the boundary pixels of the various classes in land cover use and classification. Several important points relating to the concept of the spatial connection (Liu and Shi, 2006) are as follows:

(i) **(Support of $A$):** $\text{Support}(A)$ or $\text{Supp}(A)$ is equal to the set $\{x \in X : A(x) > 0\}$. Denoted the closure of $\text{Supp}(A)$ in background topology by $\text{Supp}(A)$.

(ii) **(Supported separated):** Let $(X, I_X, \delta)$ be an I-fts and $(X, \beta)$ be its background topology space, $A, B \in I_X$. A and $B$ are called supported separated, if $\text{Supp}(A) \cap \text{Supp}(B) = \text{Supp}(A) \cap \text{Supp}(B) = \emptyset$.

(iii) **(Supported connected fuzzy set):** A is called supported connected in $(X, I_X, \delta, \beta)$, if there does not exist supported separated $C, D \in I_X \setminus \{0\}$ such that $A = C \lor D$ and $\text{Supp}(A) = \text{Supp}(C) \lor \text{Supp}(D)$.

(iv) Let $(X, I_X, \delta, \beta)$ and $(Y, I_Y, \mu, \gamma)$ be two I-fuzzy spaces, $f : X \rightarrow Y$ an ordinary continuous mapping. If $A \in I_X$ is a supported connected fuzzy set, then $f^{-1}(A)$ is also supported connected fuzzy set.

In this study, an attempt is made to introduce the fuzzy topology classification method into the three classifiers, included, MLC, MIN and MAH. The aim is to find the best combination of these classifiers leading to a higher accuracy of classification.
Table 1
The classification methods discussed in this paper.

<table>
<thead>
<tr>
<th>Classification methods</th>
<th>MLC</th>
<th>MIND</th>
<th>MAH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single classifier classification methods</td>
<td>MLC</td>
<td>MIND</td>
<td>MAH</td>
</tr>
<tr>
<td>Fuzzy single classifier classification methods</td>
<td>Fuzzy MLC</td>
<td>Fuzzy MIND</td>
<td>Fuzzy MAH</td>
</tr>
<tr>
<td>Simple average MCS</td>
<td>Simple MLC + MIND</td>
<td>Simple MLC + MAH</td>
<td>Simple MLC + MIND + MAH</td>
</tr>
<tr>
<td>Fuzzy topological simple average MCS</td>
<td>Fuzzy Simple MLC + MIND</td>
<td>Fuzzy Simple MLC + MAH</td>
<td>Fuzzy Simple MLC + MIND + MAH</td>
</tr>
<tr>
<td>Eigen-value MCS</td>
<td>EIG MLC + MIND</td>
<td>EIG MLC + MAH</td>
<td>EIG MIND + MAH</td>
</tr>
<tr>
<td>Fuzzy topological eigen-value MCS</td>
<td>Fuzzy EIG MLC + MIND</td>
<td>Fuzzy EIG MLC + MAH</td>
<td>Fuzzy EIG MIND + MAH</td>
</tr>
</tbody>
</table>

4. The logic flow of classification methods

In this section, the logic flow of both the non-fuzzy classification scheme and the fuzzy classification is described. The three classifiers, MLC, MIND and MAH, were used by many researchers for MCS (Huenupán et al., 2008; Fumera and Roli, 2005; Mesev et al., 2001; Steele, 2000; Xu et al., 1992; Kittler et al., 1998). The classification methods discussed in this paper are listed in Table 1.

As indicated above, the major objective of combining sets of classifiers is to achieve a higher classification accuracy than that of the best individual classifier in the set. Those classification methods are then chosen for use in the fuzzy method, described in Section 3. Fig. 6 shows the logic flow of the fuzzy topological methods. In addition, in order to choose the most appropriate MCS methods in terms of accuracy, we have to test the combining set of classifiers by the simple average method and the eigen-values method. The logic flows of those classification methods show in Fig. 7.

4.1. Single classifier classification method and fuzzy single classifier classification method

The mean, variance and variance–covariance matrix are first defined for each class in the single classifier classification method. Hence, the probability density function (pdf) for each pixel is

Fig. 6. The logic flow of the fuzzy topological methods.

Fig. 7. The logic flows of the three categories of the classification methods.
obtained based on the classifier itself. The class of each pixel is then determined. In Fig. 7, the first column shows the steps of the single classifier classification method.

For the fuzzy single classifier classification method, the mean, variance and variance–covariance matrix for each class, are first determined. Hence, the probability density function for each pixel is obtained based on the classifier itself. The interior and boundary of the classes are then determined. Finally, the boundary pixels are identified, making use of the interior pixel connectivity analysis (Shi et al., 2009). Fig. 7, the second column shows the steps of the fuzzy single classifier classification method.

### 4.2. Simple average MCS and fuzzy simple average MCS

For the simple average MCS, the mean, variance and variance–covariance matrix for each class are first estimated. For each pixel, the probability density matrix is then computed with each classifier given the same weight. Finally, the class of this pixel is determined. In Fig. 7, the third column shows the steps of the eigen-value MCS.

For the fuzzy simple average MCS, the mean, variance and variance–covariance matrix for each class are first estimated. Then, for each pixel, the probability density matrix is computed and each classifier is given the same weight. The interior and boundary of classes is then determined. Finally, the boundary pixels are determined, making use of the interior pixel connectivity analysis. In Fig. 7, the fourth column shows the steps of the fuzzy topology and simple average MCS.

### 4.3. Eigen-value MCS and fuzzy eigen-value MCS

For the eigen-value MCS, the mean, variance and variance–covariance matrix for each class are first estimated. The eigen-values of the probability density matrix for each pixel are then computed and the classifier is weighted based on the quantity of the eigen-values. Finally, the class for each pixel is determined. Fig. 7, the fifth column shows the steps of the eigen-value-based MCS.

For the fuzzy eigen-value MCS, the mean, variance and variance–covariance matrix for each class are first estimated. The

---

**Table 2**

<table>
<thead>
<tr>
<th>Single classifier classification method vs fuzzy single classifier classification method.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa value</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Single classifier classification method</td>
</tr>
<tr>
<td>MLC 0.7274</td>
</tr>
<tr>
<td>MIND 0.7191</td>
</tr>
<tr>
<td>MAH 0.7619</td>
</tr>
<tr>
<td>Fuzzy single classifier classification method</td>
</tr>
<tr>
<td>Fuzzy MLC (alpha = 0.74)</td>
</tr>
<tr>
<td>Fuzzy MIND (alpha = 0.86)</td>
</tr>
<tr>
<td>Fuzzy MAH (alpha = 0.72)</td>
</tr>
</tbody>
</table>

---

**Table 3**

<table>
<thead>
<tr>
<th>Simple average MCS vs fuzzy simple average MCS.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa value</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Simple average MCS</td>
</tr>
<tr>
<td>MLC + MIND 0.7545</td>
</tr>
<tr>
<td>MLC + MAH 0.7463</td>
</tr>
<tr>
<td>MIND + MAH 0.7703</td>
</tr>
<tr>
<td>MLC + MIND + MAH 0.7582</td>
</tr>
<tr>
<td>Fuzzy simple average MCS</td>
</tr>
<tr>
<td>Fuzzy MLC + MIND (alpha = 0.66)</td>
</tr>
<tr>
<td>Fuzzy MLC + MAH (alpha = 0.72)</td>
</tr>
<tr>
<td>Fuzzy MIND + MAH (alpha = 0.66)</td>
</tr>
<tr>
<td>Fuzzy MLC + MIND + MAH (alpha = 0.72)</td>
</tr>
</tbody>
</table>

---

**Table 4**

<table>
<thead>
<tr>
<th>Eigen-value MCS vs fuzzy eigen-value MCS.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa value</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Eigen-value MCS</td>
</tr>
<tr>
<td>EIG MLC + MIND 0.7568</td>
</tr>
<tr>
<td>EIG MLC + MAH 0.7487</td>
</tr>
<tr>
<td>EIG MIND + MAH 0.7714</td>
</tr>
<tr>
<td>EIG MLC + MIND + MAH 0.7664</td>
</tr>
<tr>
<td>Fuzzy eigen-value MCS</td>
</tr>
<tr>
<td>Fuzzy EIG MLC + MIND (alpha = 0.76)</td>
</tr>
<tr>
<td>Fuzzy EIG MLC + MAH (alpha = 0.72)</td>
</tr>
<tr>
<td>Fuzzy EIG MIND + MAH (alpha = 0.76)</td>
</tr>
<tr>
<td>Fuzzy EIG MLC + MIND + MAH (alpha = 0.76)</td>
</tr>
</tbody>
</table>
Fig. 9. The classification results of various classification methods.
eigen-values of the probability density matrix for each pixel are then computed and each classifier is weighted, based on the quantity of the eigen-values. The interior and boundary classes are then determined. Finally, the boundary pixels, following the connectivity analysis, are then combined with the interior pixels. Fig. 7, the sixth column shows the steps of the fuzzy topology and eigen-value-based MCS.

5. The experimental studies

Several experiments have been conducted to test the performance of the classification methods listed in Table 1. The test area was located in the south-west of Xuzhou in China. There were four categories of land cover class: building, woodland, water and farmland, however the classification methods can be used for land cover...
classification with the class number that is larger than four. In the test area, some land cover objects were with crisp and others with natural boundaries. A Landsat satellite image taken in 2000 was used as the test data, and in fact the methods are not limited to using Landsat images. A ground survey was also carried out for verifying some of the reference data that are used for accuracy assessment and kappa statistics analysis. Landsat images, with seven spectral bands, were taken, and bands 1, 2, 3, 4, 5 and 7 were used for image classification. In this experimental study, the classification methods were conducted using MATLAB 7.1 to assign each pixel, on the image, to a specific class. Fig. 8 is the original satellite image taken in 2000 which was used for the classification test. Fig. 9 shows the results of the experiment using different classification methods. We suggest the readers to read the classification result maps together with the corresponding statistical results, listed in Tables 2–4.

6. The optimum threshold value, kappa and accuracy analysis

In the field of remote sensing applications, the medium resolution of the satellite image is one of the sources causing the problem of mixed pixels. Mixed pixels problems may be due to the boundary of spatial objects or small spatial objects, etc. The introduction of fuzzy topology and its connectivity theory between classes, we can solve many mixed pixels problems which included the problems of boundary, small spatial objects, etc. Theoretically, fuzzy topology (Liu and Shi, 2006) is induced by the value of \( \alpha \) (see formula (1)) is the intercorrelation coefficient for two classes. Consequently, as indicated above, each class is split into three parts: an interior, a boundary and an exterior. The interior of class \( c_i \) is denoted by \( (A_{ci})_{\alpha} \), the boundary of \( c_i \) is defined as \( \partial A_{ci} = (A_{ci})_{1-\alpha} \land (A_{ci})_{\alpha} \) and the exterior is not studied in this classification scheme. In finding the optimum threshold value with respect to the accuracy in the classification scheme, the graph of the threshold value against overall accuracy value has to be plotted for each classification method. Thus, the graph of the threshold value against the kappa value was plotted to find the optimum threshold value. The graphs were divided into three groups, single classifiers classification method, simple average MCS and eigenvalue MCS. In each group, the best classification methods can be compared and the optimum threshold value estimated. The experimental results show that the fuzzy topological method always produces higher accuracy and a better kappa value than the simple average MCS. From the graphs of several simple average methods, the optimum threshold and kappa values and overall accuracy, occurs at about the threshold value for the single classifiers, MLC, MIND and MAH. From the curve of MLC it can be seen that the global optimum threshold value, kappa value and overall accuracy, occur for threshold values in the range 0.66–0.78 (see Table 3). In addition, it is noted that, the fuzzy single classifier always produces more accurate results than the non-fuzzy single classifier classification method. The MLC fuzzy single classifier, therefore, is a better classification method in terms of classification accuracy and kappa value.

Fig. 10(a) is the graph on the threshold value against the kappa value for the single classifiers, including MLC, MIND and MAH. While Fig. 10(b) shows the threshold value against overall accuracy for the single classifiers, MLC, MIND and MAH. From the curve of MLC it can be seen that the optimum threshold value, kappa value and overall accuracy, occur at about the threshold value of 0.74 and 0.84, respectively. From the curve of MIND it can be seen that the optimum threshold value, kappa value and overall accuracy, both occur at 0.86 and 0.86, respectively. From the curve of MAH, we can see that the optimum threshold value, kappa value and overall accuracy, occur at 0.72 and 0.72, respectively (see Table 2). In addition, it is noted that, the fuzzy single classifier always produces more accurate results than the non-fuzzy single classifier classification method. The MLC fuzzy single classifier, therefore, is a better classification method in terms of classification accuracy and kappa value.
As expected, the fuzzy topological classification for simple average MCS is always better than the simple average MCS in terms of classification accuracy or kappa value. In addition, the MLC + MAH fuzzy simple average MCS is the best method when compared to simple average MCS.

Fig. 12(a) shows the threshold value against kappa values for eigen-values MCS, included MLC + MIND, MLC + MAH, MIND + MAH and MLC + MIND + MAH. Fig. 12(b) shows the threshold value against overall accuracy for eigen-values MCS. From the graphs of several simple average methods, the optimum threshold and kappa value and overall accuracy, occur in the threshold range 0.72–0.76 (see Table 4). As expected, the fuzzy eigen-values MCS is always better than the eigen-values MCS as far as classification accuracy and kappa value are concerned. In addition, the MLC + MAH fuzzy eigen-values MCS is the best method when compared to eigen-values MCS.

After comparing the classification results in visual sense (see Fig. 9), the kappa value and overall accuracy graphs, it is not difficult to come with the following three conclusions:

(i) The fuzzy topological method is always better than the ordinary method. For example, in Fig. 13(a), with the method of MLC + MIND + MAH eigen-value MCS, we note that many pixels (those as very small dots within the ellipse) are wrongly classified but they are correctly classified in Fig. 13(b), with the method of MLC + MIND + MAH fuzzy eigen-value MCS.

(ii) The eigen-value MCS method is better than the non-fuzzy method in terms of kappa value and overall accuracy.

For example, the simple MIND + MAH have the kappa value and the overall accuracy equal to 0.7703 and 0.8657, respectively. On the other, the eigen-value MIND + MAH gives better kappa value and overall accuracy, they are 0.7714 and 0.8664, respectively.

(iii) MLC + MIND + MAH fuzzy eigen-value MCS is the best method as this method attains highest kappa value equal to 0.8089 when alpha equal to 0.76; and overall accuracy equal to 0.8952 when alpha equal to 0.76 (see Table 4). Fig. 13(a) and (b) shows several improvement visually.

Fig. 11. (a) Threshold value against kappa value for simple average MCS. (b) Threshold value against the overall accuracy for simple average.

Fig. 12. (a) Threshold value against kappa value for eigen-values MCS. (b) Threshold value against Overall accuracy for eigen-values MCS.

Fig. 13. (a) The ordinary method with many misclassifications – in very small dots. (b) The fuzzy topological result that correct misclassifications which are in small dots.
7. Conclusions

The multiple classifier system (MCS) method is an effective automatic classification method for use in remote sensing analysis. Combined with the induced fuzzy topology concept, a method developed in previous studies (Liu and Shi, 2006), enables an MCS decomposition of image classes. This fuzzy topological MCS provides a new classification approach for obtaining a higher classification accuracy. That is, the fuzzy topological space through class decomposition to re-classify the mixed pixels. Which the mixed pixels pixels cause by the boundary of spatial objects or small spatial objects can be solved. A four classes, building, woodland, water and farmland, land cover classification experiment enabling the optimum classification method combination to be observed, has been demonstrated. There are several important points about this fuzzy topological classification method.

(a) The reason of choosing those four classes is that they have a high degree of connectivity so that the connection theory in fuzzy topological space was applied to combine the class interior and boundary.

(b) This fuzzy topological classification method not only can be applied on Landsat image, but also on other images.

(c) Moreover, the fuzzy topological classification method has been applied on the four classes land cover classification experiment in this paper. But it can be extended to large numbers of classes' classification.

(d) Other than land cover classification, the fuzzy topological classification method can be easily extended to other pattern recognition applications.

The performance of the various methods, which include the single classifier classification method, simple average MCS, eigen-value MCS, fuzzy single classifier classification method, fuzzy simple average MCS and fuzzy eigen-value MCS, have been evaluated by several experiments, by means of a newly designed procedure, with the MAH being found to be the best individual classifier for the application of the fuzzy topological method. Thus, a significant improvement in the kappa value and overall accuracy is achieved once the concept of connection in fuzzy topology is introduced. Further found is that the MLC + MIND + MAH fuzzy eigen-value MCS obtained the highest kappa value and overall accuracy. This is because the weights of the classifiers are based on the eigen-values. That is, if a large classifier value, is assigned to a large eigen-value, can enhance the effect of that classifier, naturally. Similarly, a small classifier value, is assigned to a small eigen-value, can reduce the effect of that classifier is naturally enhanced. Note also that, the MLC, the MIND and the MAH are complementary to each other in this experiment.

In general, the fuzzy method based classifications are better than those of the non-fuzzy methods. This is due to the introduction, into the classification, of the topological space concept enabling the concept of fuzzy topology connection to be applied to enhance the classification process.

In overall summary, the contributions of this paper include: (a) the development of a classification method list, includes the single classifier classification method, simple average MCS, eigen-value MCS, fuzzy single classifier classification method, fuzzy simple average MCS and fuzzy eigen-value MCS; (b) the newly developed methods have been shown to provide better classification accuracy and kappa values than the non-fuzzy methods; (c) the experiment based on real satellite data were successful in identifying the best the best combination of fuzzy classification methods.

Further development to this study includes the extension of the fuzzy topological method to other possible classification methods.

Acknowledgments

The work presented in this paper is support by Ministry of Science and Technology (Project No. 2006BAJ05A08), The Hong Kong Polytechnic University (Project No. G-YX0P, G-YF24, G-YG66, 1-ZV4F) and Hong Kong RGC General Research Fund (Project No. 5276/08E).

References


