INCREASING CLASSIFICATION QUALITY BY USING FUZZY LOGIC

DROJ Gabriela,
University of Oradea, e-mail: g_abyg@yahoo.com

ABSTRACT
Developed initially for the storage of the graphical information mainly maps into digital format, in the last decades geographical information systems have become one of the main supporting instruments for the decision making in the various fields. This development led to the necessity of reducing processing time by using more advanced applied mathematical and computer science knowledge. One of these mathematical theories, which give the possibility of enhancing spatial data management with the modelling of uncertainty, is fuzzy logic. This paper will present possibilities for increasing the quality of image classification by using fuzzy logic.

Keywords: GIS, image classification, fuzzy logic

INTRODUCTION
Developed initially for the storage of the graphical information mainly maps into digital format, in the last decades geographical information systems have become one of the main supporting instruments for the decision making in the various fields. Geographical information systems are used for decision support tool in fields like urban planning, investment planning, economic development, resource management, impact studies and in critical systems. The evolution of GIS led to the necessity of faster and better results, requirement fulfilled with the increase usage of advanced and applied mathematical and computer science knowledge like spatial databases or complex algorithms from computational geometry, artificial intelligence and fuzzy logic. This paper will present possibilities for increasing the quality of image classification by using fuzzy set theory.

MATERIALS AND METHODS
The fuzzy set theory gives the possibility to use computers in representing and processing uncertain data or incomplete rules, and information described in a common language and for which no certainty degree can be determine.

In fuzzy logic geospatial data or layers are evaluated based on qualitative information and not for quantitative ones. In this way the fuzzy logic theory gives the possibility of enhancing spatial data management with the modelling of uncertainty. Fuzzy logic was implemented successfully in various GIS processes [2] like:

- Data collection analysis and processing remote sensing data for classification algorithms and object recognition [1];
- Spatial analysis - processing qualitative data, in order to defining relationships between uncertain geospatial objects;
- Complex operations based on genetic algorithms or artificial intelligence; like object recognition from airborne images [2].

In many cases fuzzy concepts are more suitable for modeling information like land cover or land use, because there is no exact threshold between different categories of similar information like
vegetation. In order to define boundaries between such areas, clear polygon boundaries are difficult to be determined. In such cases a widely used procedure is to allocate all grid cells to classes based on fuzzy membership than 0/1 membership. Membership functions are mathematical function which defines the degree of an element’s membership in a fuzzy set.

- **The sigmoidal** (“s - shaped”) membership function is perhaps the most commonly used function in fuzzy set theory it is produced by a combination of linear and \( \cos^2() \) functions [3].
- **J-shaped functions**, which are rather like the sigmoidal membership function but with the rounded top sliced off flat over some distance [3].
- **Linear functions**, which are like the J-shaped function but with linear sides, like the slope of a pitched roof, and are thus simple to calculate and have a fixed and well-defined extent, and
- **User-defined functions**, which are self-explanatory.

In most applications membership functions are symmetric, although monotonic increasing or decreasing options are provided in Idrisi [3].

![Fig. 1. Sigmoidal Membership Functions](image1)

![Fig. 2. J-shaped Membership Functions](image2)

![Fig. 3. Linear Membership Functions](image3)

Fuzzy set membership is used in order to obtain the standardization of criteria. Exactly which function should be used will depend on the understanding of the relationship between the criterion and the decision set, and on the availability of information to infer fuzzy membership. In most cases, either the sigmoidal or linear functions will be sufficient. Image classification represents the processing technique, through which the data represented in the image (pixels) is grouped in a certain number of classes. The classification methods are divided in two big categories: supervised clusterization and unsupervised clusterization. Supervised clusterization is the process of classification assisted by the user. In the process of supervised classification we can distinguish three stages: learning, assigning and testing. The training stage consists in collection of references, information and data about the image which should be classified, which will be used for defining a vectorial file, for each cluster being defined a layer within the file. The result of the learning stage is represented by the signature type files which are characterizing the information which should be classified.
In the second stage: of assigning, the pixels from the image are allocated in the basis of static results in accordance with the signatures files. The final stage is testing of the results by comparing the classified image with the previous data. The pixels which were classified in different clusters are stored in the error matrix and later being analyzed by the user. The fuzzy logic can be used in the creation of the signatures files and also in the process of the classification itself, but is not necessary to be used in both stages [1, 3, 4, 5, 6, 7].

The signatures files based on this training site can be created using unsupervised classification methods (like ISODATA), classical methods (like Maximum Likehood or Minim Distance) or by computing a fuzzy matrix filled with the values indicated by the membership grade of each training site [3]. The membership degree is computed by using fuzzy membership functions like: sigmoidal, linear or j-shaped presented before. The evaluation on the training can be done using statistical methods: minimum, maximum, mean, and standard deviation for each band independent and covariance matrix for all the three bands [2]. The most relevant signature file evaluation is creating an error matrix as a matrix of percentages based on pixel counts that allows us to see how many pixels in each training sample were assigned to each class [3, 8].

The second step of the supervised classification can also be processed with traditional methods (Minimum distance, Mahalanobis distance, Maximum Likehood so on) or by using fuzzy membership functions or genetic algorithm [2]. This stage cannot be evaluated alone, it can be evaluated just the final result of the supervised classification [8]. In order to optimize the accuracy of the results we consider that is appropriate to use fuzzy logic in both stages of supervised classification. In the next chapter we are testing these hypotheses [2].

Independently of the classification method used, in order to increase the quality of the results we propose usage of the fuzzy convolution operation. Fuzzy convolution allow to create a single classification band by calculating the total weighted inverse distance of all the classes first and assigning the center pixel the class with the largest total inverse distance summed over the entire set of fuzzy classification bands [9]. This has the effect of creating a context-based classification to reduce the noise of the classification [3, 4, 9]. Classes with a very small distance value remain unchanged while classes with higher distance values may change to a neighboring value if there are a sufficient number of neighboring pixels with class values and small corresponding distance values [3, 4].

RESULTS AND DISCUSSIONS

Input data. For the procedures of image classification was used an orthorectified airborne image from the upper hills of Oradea municipality. This image contains three channels recorded in three bands: the first band for green, the second for red and the third for blue. In the figure 4, we present this image [2].

Standardization criteria In order to analyses the standardization criteria for the build area the minimum distortion of the border of it was studied the possibility of using different membership functions, which function should be used depend relationship between the criteria of classification [8].
In the first experiment we used a sigmoidal monotonic decreasing membership function. The result is represented in the image presented in Figure 5a). The test was repeated for linear monotonic decreasing membership function but the result was identical.

In the second case we used a j-shaped monotonic increasing membership function. The result is represented in the image presented in Figure 5b).

**Supervised classification.** Training is the first stage of a supervised classification. In this step the user must define training areas for each class interactively on the displayed image. The areas may be specified both on polygon and on pixel basis. The three classes of information defined are: green areas, houses and streets. The signature files were created using a fuzzy membership function based on the j-shaped membership function presented before, in this case the accuracy of the signature files are acceptable, over 90 % (table below) [2].
The classification of the input images, with the signature files defined before, was done by using a classifying method which combined the Maximum Likelihood method with a fuzzy membership function (linear). The output images are represented in the figure 7, each image represent a cluster in the following order houses, green areas, streets [2, 8].

![Fig. 7. Supervised classification](image)

Classification results can be optimized through fuzzy convolution. The scope of this procedure is to combine the previously created bands in one band by interpolation. The interpolation method used is the weight inverse distance. Applying the fuzzy convolution on the results, obtained after the second case of supervised classification, we obtained the image below [2].

![Fig. 8. Fuzzy convolution](image)

**CONCLUSIONS**

The mathematical set theory used in GIS software enforce an artificial precision sometimes on vague information about the real world, fuzzy logic propose a new way to handle the uncertainty of the real world.

Fuzzy procedures could increase the quality of raster image classification and object recognition.

<table>
<thead>
<tr>
<th>Data</th>
<th>streets</th>
<th>houses</th>
<th>green areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>streets</td>
<td>95.45</td>
<td>4.87</td>
<td>1.27</td>
</tr>
<tr>
<td>houses</td>
<td>2.73</td>
<td>93.42</td>
<td>2.83</td>
</tr>
<tr>
<td>green areas</td>
<td>1.82</td>
<td>1.71</td>
<td>95.89</td>
</tr>
</tbody>
</table>
Fuzzy logic offers instruments for both types of classification. The unsupervised fuzzy-based classification allows clustering of data, where no initial information is known, but the supervised classification is offering higher quality.

The experiments revealed that by using fuzzy-based algorithm in all stages of the supervised classification the quality of the result is increased [2].

We demonstrated through the experiments presented in this paper that fuzzy logic could be used to increase the quality of classification. These represent a follow-up of previous experiments described in other works [2, 8] when we demonstrated that the fuzzy set theory could be used as a very powerful tool in designing and implementing algorithms for geographical informational systems and also to support spatial decision-making process.

REFERENCES

2. DROJ, GABRIELA (2007), The Applicability Of Fuzzy Theory In Remote Sensing Image Classification, STUDIA University Babes-Bolyai, INFORMATICA, Volume LII, Number 1.
8. DROJ, GABRIELA (2011), The Applicability of Fuzzy Theory In Spatial Data Analyses, Geoprevi.