2010 International Conference on Distributed Framework for Multimedia Applications (DFmA)

Color Pixel Classification Using Genetic Fuzzy System: Case Study on Earth Surface Classification

Budi Darma Setiawan[1], Subanar[2]
Department of Computer Science
Gadjah Mada University
Sekip Utara Bulaksumur, Yogyakarta 55281, Indonesia
Telp/Fax: 62 274 555133
[1]s.budidarma@gmail.com
[2]subanar@yahoo.com

***Abstract* -- From satellite image, people can see any objects on the earth surface like houses, streets, lands, vegetation, and water. To classify those objects, observers have to be able to distinguish objects in the image. One of the simplest methods is by analyzing pixel color. In this case, the fuzzy classification system is chosen since there are some overlapping in the pixel color characterized for certain objects. Problem of this system comes when there are no available rules for describing the classification. Therefore, genetic fuzzy system is used for creating the rules. This training process is divided in to two steps which are, learning process to create the group of rules, and tuning process to optimize the fuzzy membership function. The result is measured by CCR (Correct Classification Rate). During the training process, the CCR values are increased after the tuning process is done. The highest CCR value recorded for training process is 82,49%.Value of execution time depends on the number of rules that are used.**

***Keyword* – Fuzzy Classification System, Genetic Algorithm, Genetic Fuzzy System, Automatic Design of Fuzzy system**

I. INTRODUCTION

From satellite image, people can see any objects on the earth surface like houses or building, streets, lands, vegetations, and water. To classify these objects, observers have to be able to distinguish objects in the image. The simplest way to do is by analyzing color pixel of the image.

The use of color in image processing is motivated by two principal factors. First, color is powerful descriptor that often simplifies object identification and extraction from scene, and second, human can discern thousands of color shades and intensities compare to about only two dozen shades of gray [1]. Digital images consist of pixels and each pixel contains color information. This color information is represented in some color models, one of them is RGB (Red, Green, and Blue) model. Images used in this research have 24 bit of color depth, it is mean that every color component (Red, Green, and Blue), has 8 bit color depth or can be represented by value 0 to 255.

In the case of earth surface classification, some objects have resembled color to other object. These are shown by some overlapping pixel color for certain objects in different classes and that is why fuzzy concept and fuzzy classification system suitable for this problem.

The main component of fuzzy system is a knowledge consisting fuzzy IF-THEN rules [2]. The problem comes when we want to build this system without knowing the rules. Some methods have been proposed to create this system [2]. One of the methods is by using genetic algorithm [3]-[6].

This research describing how to use genetic algorithm to creates the rules for classifying the pixel color in the image taken from Google Earth, through data sample training.

For the first step, in this research, some training data are created by supervised classification through random pixels. These pixels are taken from random images in *Google Earth*. After that, The Genetic algorithm will search the best combination of the rules that mostly match the training data. Lastly, Genetic algorithm is used again in tuning the membership function to maximize the accuracy.

A research for classifying earth surface image using fuzzy classification has been done before [8]. That research uses signature statistic from supervised classification to create rule and membership function, meanwhile our research proposes to use genetic algorithm to create fuzzy rule to classify earth surface.

Section 2 is filled with review of Fuzzy classification system. Section 3 reviews genetic algorithm and genetic fuzzy system. Section 4 shows the design of fuzzy classification system. Section 5 explains the construction of fuzzy classification system for the problem. And section 6 and 7 shows the result and conclusion of this research respectively.

II. FUZZY CLASSIFICATION SYSTEM

A Non-Fuzzy classification system normally assumes that the pattern x belongs to only a particular class according to given criterion and can be used to classify classes with well separable, well defined and distinct boundaries. For problem with overlapping region classes, fuzzy classifier is more applicable [9].

Fuzzy classification system is Rule Based Fuzzy System designed to perform classification task [10]. Example of the rule that commonly used is:

Rp: IF x_1 is A_1 and x_2 is A_2 and … and xn is A_n Then
Class is C_m

Where R_p denotes the p^{th} rule, x_n denotes nth attribute or feature taken for classification, A_n is the nth linguistic value for the attribute, and C_m denotes the class which the object is classified.

In [10], there is explanation on steps how to use the set of rules for classification process. First let $e_k = (a_{k1}, a_{k2}, …., a_{kn})$ be a pattern to be classified and $\{R_1,R_2,…, R_p\}$ is the set of S rules, each with n antecedents. Let the $A_i(a_{ki})$, i=1,…,n is the membership degree of attribute value aki in ith fuzzy set of fuzzy rule R_p. And the steps to get the conclusion are:

1. Calculates the compatibility degree between pattern ek and each rule R_p for p=1,…,S
   $\mathrm{Comp}(R_p, e_k) =$
   t-norm $(A_1(a_{k1}), A_2(a_{k2}), …, A_n(a_{kn}))$
2. Finds the rule $R_p$max with the highest compatibility with the pattern
   $R_pmax = \mathrm{Max}(\mathrm{Comp}(R_p, e_k))$
   p = 1,2,…,S
3. Assign class C_j to the pattern e_k, where C_j is the class given by rule R_pmax from previous step

III. GENETIC ALGORITHM AND GENETIC FUZZY SYSTEM

Genetic Algorithm (GA) is an algorithm adapted from genetic behavior of the living things. This genetic thing is represented by chromosomes which carry characteristic of an individual. Chromosomes are evolved in each generation to create new individual (offspring) with new characteristic by mating process. In the mate process, there is a process that exchanges genes among chromosomes (recombination crossover) and process that alter the gene value (mutation).

In Genetics algorithm, all features value that characterize the problem are coded into the chromosomes. Then, those chromosomes evolve generation to generation, until one of the chromosome (individual) meet the solution of the problem. Fitness value is given to each chromosome to evaluate which chromosome is the fittest to the solution. The main steps in genetic algorithms is:

1. Initialize the population
2. Evaluate each individual
3. Repeat until terminated condition
4. Select parents to mate
5. Recombination
6. Mutation
7. Evaluation
8. Replace old population
9. End Repeat

Genetic Algorithm can be used for resolving some problem like routing problem, scheduling problem, and optimization problem [11]. It also can be used in order to design and optimize fuzzy system. This kind of combination of genetic algorithm and fuzzy system is called genetic fuzzy system [5].

Genetic fuzzy system is designed to automatically create the fuzzy system which containing two main parts, Rule Base (Set of rules) and Data base (attribute partition of membership function). According to those two different parts, genetic fuzzy systems are divided in two main steps, genetic learning process and genetic tuning process [6]. Genetic learning process is the process to automatically generate the rule for the fuzzy system, and genetic tuning process is the process to optimize the system performance.

There also some approaches to perform genetic learning process, they are: Michigan approach, Pittsburgh approach, and Iterative [5], [6]. In Michigan approach, each chromosome codes the individual rules, meanwhile in Pittsburg, each chromosome codes a rule base (all rule). Iterative approach is derived from Michigan, which iteratively put the rule to the rule base.

Each method has their own advantage and disadvantage [3]. This research chose Pittsburgh approach because it code and analyze all rules at once as a rule base, instead of doing it one by one.

IV. DESIGN OF FUZZY CLASSIFICATION SYSTEM

This research describes the process of making the fuzzy classification system to classify pixels of earth surface images to some classes. The classes are: buildings (1), streets (2), lands (3), water (4), and vegetation (5).

The feature taken from the pixels for classification attribute is the pixel's color. This pixel's color is represented in RGB format. Each R, G, and B has value 0 to 255 and has their own definition of membership function.

Initially, each feature is divided in 5, 6, and 7 membership function partition, as shown in figure 1, where VD = very dark, D = dark, MD = medium dark, M = medium, MB = medium bright, B = bright, VB = very bright. Those three division kind of membership function are tested to get which is the most efficient among them.

These initial membership function are used in genetic learning process to learn fuzzy rule, and then rearranged in genetic tuning process to increase classification accuracy. Genetic learning and genetic tuning process will be describe in next section.

Example of rules generated from the genetic learning process are:

IF r is VD and g is VD and b is VD THEN Class is 4 (water)
IF r is VD and g is VD and b is D THEN Class is 2 (land)
IF r is VD and g is VD and b is M THEN Class is 1 (building)

Where r, g, b is intensity (0 to 255) of Red, Green, and Blue respectively.

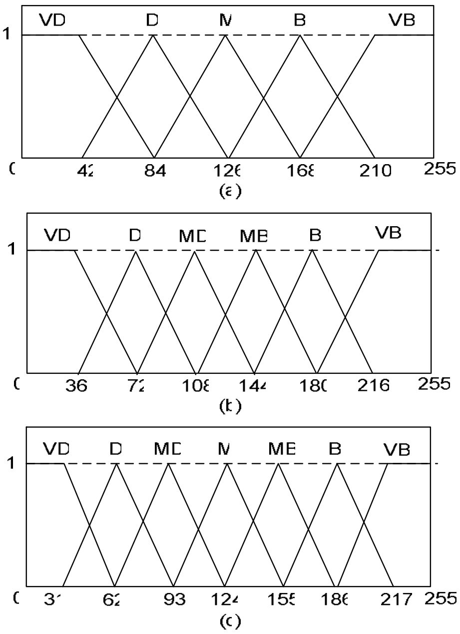

Fig.1 Initialization of membership function partition:
(a) 5, (b) 6, and (c) 7 partitions

V. CONSTRUCTION OF FUZZY SYSTEM

A. *Data Training from supervised classification*

Training data used to build this fuzzy classification system are images with 10 x 10 pixels. These images are taken randomly from Google Earth. All pixels contained in each training image represent color of one class.

In figure 2, there is an example of the image taken from Google Earth and the yellow spot is the place where 10 x 10 training images are taken.

Fig. 2 Example of training images

B. *Genetic Learning Process*

Genetic learning process is the process to learn or create a rule base for the fuzzy system. This process uses genetic algorithm with 100 chromosomes in each generation, with crossover and mutation ratio are among 0 to 1. Genetic process (selection, mutation, and crossover) are attempted to be repeated for 100, 250, 500, 1000, 1250, and 1500 generation. The variation number of generation is considered since we do not know optimum capability of fuzzy system generated by this method.

Since using Pittsburg approach, a chromosome is now consist of combination from all possible rules. To avoid rule duplication, position of genes denotes the combination of linguistic value from membership function for each feature (all variation of antecedent combination), and the genes values denote the class number (consequent part of the rue). The genes values are coded using integer with range from 0 to 5 (0 for not used rule, and 1 to 5 for class number).

Because gene position denotes the different combination of antecedents, and if each features RGB is divided into 5 membership function, so there are 5x5x5 = 125 genes in one chromosome. Figure 3 shows the genetic representation.

Gen:	1	2	3	4		……		125
	0	2	1	0		…….		5

Fig. 3 Example of genetic representation in learning process

Chromosome in figure 3 means:
Gen1: IF r is **VD** and g is **VD** and b is **VD** THEN Class is 0 (rule is not used)
Gen2: IF r is **VD** and g is **VD** and b is **D** THEN Class is 2 (Lands)
Gen3: IF r is **VD** and g is **VD** and b is **M** THEN Class is 1 (Buildings)
……
Gen6: IF r is **VD** and g is **D** and b is **VD** THEN …………..
Gen7: IF r is **VD** and g is **D** and b is **D** THEN …………..
Gen8: IF r is **VD** and g is **D** and b is **M** THEN ………….
……
Gen123: IF r is **VB** and g is **VB** and b is **M** THEN ………….
Gen124: IF r is **VB** and g is **VB** and b is **B** THEN ………….
Gen125: IF r is **VB** and g is **VB** and b is **VB** THEN Class is 5 (Vegetation)

In this process, there are two things can be considered as variable to compute the fitness value. They are: (1) rate of compatibility of rule base to classify object correctly. It is called Correct Classification Rate (CCR) [10]. And (2) number of the rule generated. Since there are more then one objectives to evaluate the chromosomes, this genetic process use multi objectives genetic algorithm concept.

Chromosome's CCR is given by

$$CCR = \frac{number\ of\ true\ classification}{number\ of\ all\ samples}$$

Fitness value is determined by using both objectives. To make this process easier, objectives are set to be optimum if they are in their maximum value. For CCR, it fulfills the condition, but for number of rule, it has to be count inversely. We count the number of unused rule instead count the used rule.

Then fitness value is determined using adaptive weight sum approach [11]. It is given by:

$$z(x) = \sum_{k=1}^{2} w_k f_k(x)$$

where $z(x)$ is the fitness value, f_k is k^{th} objective function, and w_k is weight for k^{th} objective. Weight is given by

$$w_k = \frac{1}{z_k^{\max} - z_k^{\min}}$$

for k = 1 (CCR), 2 (number of rules)

where

$$z_k = f_k(x)$$

Roulette wheel method used for selection process. n-point crossover [7] is used for crossover operation. And random resetting is used for mutating the chromosomes [7]. Lastly the replacement process use elitism to maintain the best chromosomes.

C. Genetic Tuning Process

Since this process objective is for optimizing membership function, genetic representation contains information about membership function attribute.

Membership function used in this research is trapezoidal form, because it has distinct boundaries compared to non-zero membership function (e.g. Gaussian) and has more region with membership value equal to 1 instead of triangular form. The attribute of this trapezoid membership function shown as a, b, c, and d in figure 4.

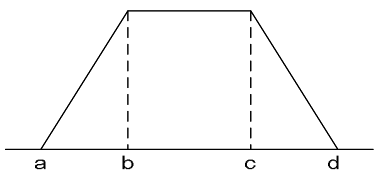

Fig. 4 Trapezoidal membership function

To optimize the membership function using genetic Algorithm, genetic representation has to include these membership function attribute. And because the classification process involve three features (R, G, and B), so the chromosomes are divided into three region. Figure 5 show the example for this genetic representation with condition that each feature (R, G, and B) is divided in to 5 partitions of membership function.

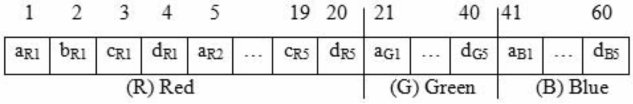

Fig. 5 Example of genetic representation for tuning process

The initial value use the value from figure 1, with value b and c are intersected.

To keep the trapezoid form, there are boundaries in genetic operation especially for mutation process [5]. The boundaries are:

$$a \in [a_l, a_r] = [a - (b-a)/2 , a + (b-a)/2]$$
$$b \in [b_l, b_r] = [b - (b-a)/2 , b + (c-b)/2]$$
$$c \in [c_l, c_r] = [c - (c-b)/2 , c + (d-c)/2]$$
$$d \in [d_l, d_r] = [d - (d-c)/2 , d + (d-c)/2]$$

The parameters on those boundaries are described in figure 6.

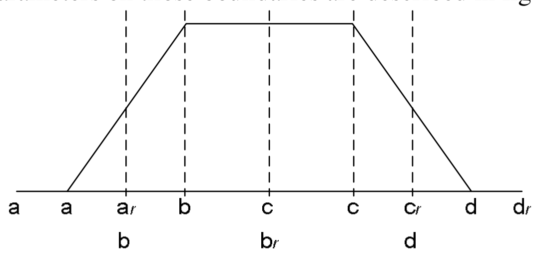

Fig. 6 Boundaries for optimizing trapezoidal membership function

Because the number of rules is fixed from genetic learning process, the fitness value for this process only use CCR value.

VI. RESULT

After some series of experiment about genetic parameter are done, the results are shown below.

First, the experiment about the number of generation. The results are shown in table 1, and graphi on figure 7 shows the pattern.

TABLE 1
RESULT FOR EXPERIMENT USING VARIATION NUMBERS OF GENERATION.

Generation Number	Learning Rule Process		CCR After Tuning
	CCR	Number of Rule	
100	0.5426	81	0.6411
250	0.6131	75	0.7630
500	0.6690	59	0.7569
1000	0.7090	61	0.8023
1250	0.7250	51	0.8136
1500	0.7200	49	0.8249

Graph in figure 7 shows that the CCR value is increasing as well as the number of generation, but getting constant at some point (1250 and 1500 generation). It shows that on that generation, the process reaches the optimum value.

The next experiment is by changing the number of partition of membership function. The result shows that since more partition used, more rules are generated, and need more generation to reach the optimum value. Table 2 shows the result.

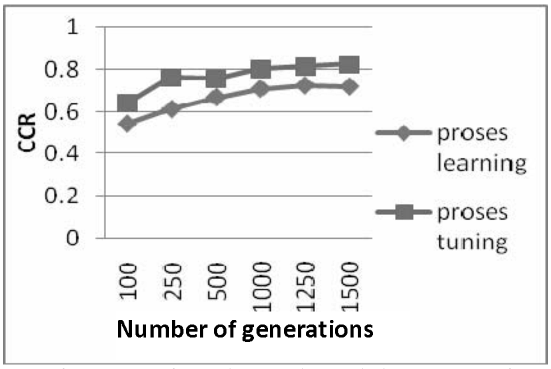

Fig. 7 Graph from result of experiment using variation numbers of generation

TABLE 2
RESULT FOR EXPERIMENT USING SOME NUMBERS OF MEMBERSHIP FUNCTION PARTITION

Numb. of Partition	Learning Rule		CCR After Tuning
	CCR	Number of Rule	
5	0.6131	75	0.7630
6	0.5253	144	0.6180
7	0.4798	242	0.6255

The drastic alteration of number of rules used for classification does effects the computation time of the system in classification process. Figure 8 shows the average of time needed for classifying all pixels in an image with resolution 15 x 15 pixels. The execution time compared with the number of rules used.

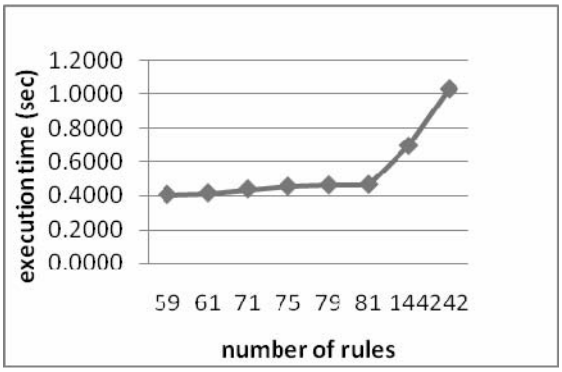

Fig. 8 Graphic of average of execution time

Table 3 shows all experiment result, and the influence of genetic tuning process in increasing the accuracy of classification process. It shows that the average increment of CCR after tuning process is 17.24%.

From the training (learning and tuning process) using 1250 Generation, crossover ratio 0.3, mutation ratio 0.2, 5 membership partition we get 51 rules and membership function shown in figure 9.

Those 51 rules and membership function are tested using some images and the result shown in figure 10. From classification result, yellow is given for buildings, red for lands, blue for water, grey for street, and green for vegetation.

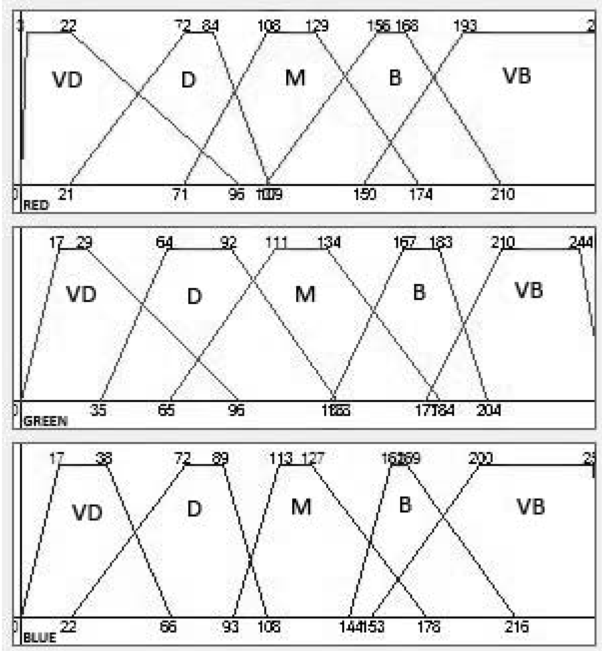

Fig. 9 Membership function graphics generated from training

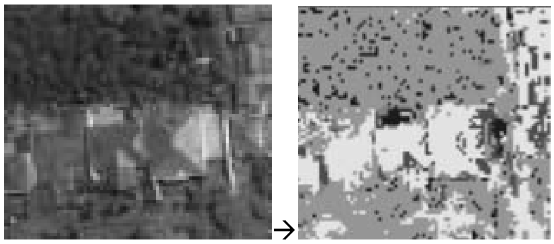

Fig. 10 Classification result

From result in figure 10 there are some little blue spots among the green, in realm however, there are no water object there. The blue spots refer to shadow. The shadows are detected as water because in training process, there is training image belongs to water class which has very dark blue color.

VII. CONCLUSION

Genetic fuzzy classification system is used to create fuzzy classification system with unknown rule base.

This problem of classifying pixel is done using genetic fuzzy system. In the process, Genetic learning process is using Pittsburg approach. The overall process of training data can create fuzzy classification system which then can classify the pixels with accuracy until 82.49%.

There are some difficulties found in this classification process. First, the shadows in the image. The shadow is always very dark or completely in black, that makes the pixels in the shadow often recognized as particular class. Second, the variety color of building's roofs. Sometime the roofs are painted with blue or green which similar to color of ocean (water) and vegetation.

TABLE 3
SOME EXPERIMENT RESULT FOR CCR VALUE AFTER LEARNING AND AFTER TUNING PROCESS

GA Parameter			**CCR**		**Difference**	**Increment (%)**
Generation	**Number of Partition**	**Crossover Ratio : Mutation Ratio**	**After Learning**	**After Tuning**		
100	5	0.3 : 0.15	0.5426	0.6411	0.0985	18.15
250	5	0.3 : 0.15	0.6131	0.7630	0.1499	24.45
250	6	0.3 : 0.15	0.5253	0.6180	0.0927	17.65
250	7	0.3 : 0.15	0.4798	0.6255	0.1457	30.37
250	5	0.1 : 0.1	0.6478	0.7370	0.0892	13.77
250	5	0.6 : 0.2	0.6608	0.7594	0.0986	14.92
500	5	0.3 : 0.15	0.6690	0.7569	0.0879	13.14
1000	5	0.3 : 0.15	0.7090	0.8023	0.0933	13.16
1250	5	0.3 : 0.15	0.7250	0.8136	0.0886	12.22
1500	5	0.3 : 0.15	0.7200	0.8249	0.1049	14.57
Increment average						17.24

REFERENCES

[1] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, 3rd ed., Prentice Hall, New Jersey, 2008.

[2] Li-Xin Wang, *A Course In Fuzzy Systems and Control*, Prentice Hall, New Jersey, 1996.

[3] Ji-Dong Li, Xue-Jie Zhang, and Yuan-Shan Chen, "Applying Expert Experience to Interpretable Fuzzy Classification System Using Genetic Algorithms", *Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)* Vol.2, pp.129-133. 2007.

[4] Xing Zong-Yi, Hou Yuan-Long, Tong Zhong-zhi and Jia Li-min, "Construction of Fuzzy Classification System Based on Multi-objective Genetic Algorithm", *Sixth International Conference on Intelligent Systems Design and Applications (ISDA'06)*, pp. 1029-1034, 2006.

[5] F.Herrera, M. Lozano, and J.L. Verdegay, "A Learning Process For Fuzzy Control Rules using Genetic Algorithms", *Technical Report DECSAI-95108, University of Granada*, Department of Computer Science and Artificial Intelligence, Granada, Spain, 1995.

[6] Frank Hoffmann, "Evolutionary Algorithms for Fuzzy Control System Design", *Proceedings of the IEEE*, 89(9):1318–1333, 2001.

[7] A.E. Eiben, and J.E. Smith, *Introduction to Evolutionary Computing*, Springer-Verlag, Berlin, 2003.

[8] I.Nedeljkovic, "Image classification based on fuzzy logic", *Proc.Geo-Imagery Bridging Continents XXth ISPRS Congress*, pp.12-23, Istanbul, 2004.

[9] S.K. Meher, "A New Fuzzy Supervised Classification Method Based on Aggregation Operator", *Third International IEEE Conference on Signal-Image Technologies and Internet-Based System*, sitis, pp.876-882, 2007.

[10] M.E. Cintra, Heloisa de A. Camargo, E.R. Hruschka Jr., dan M. do Carm Nicoletti, "Fuzzy Rule Base Generation through Genetic Algorithms and Bayesian Classifiers A Comparative Approach", *Seventh International Conference on Intelligent Systems Design and Applications* (ISDA 2007), pp.315-322, 2007.

[11] Mitsuo Gen, and R.Cheng, *Genetic Algorithms and Engineering Optimization*, John Wiley & Sons, New York, 2000.