

GENETIC PROGRAMMING AS A PREPROCESSING TOOL TO AID MULTI-TEMPORAL IMAGERY CLASSIFICATION

H. G. Momm

Department of Geology and Geological Engineering
The University of Mississippi, University, MS 38677, USA
hmomm@olemiss.edu

Greg Easson

University of Mississippi Geoinformatics Center
The University of Mississippi, University, MS 38677, USA
geasson@olemiss.edu

Dawn Wilkins

Computer Science and Information Science,
The University of Mississippi, University, MS 38677, USA
dwilkins@sunrise.cs.olemiss.edu

ABSTRACT

Classification-based applications of remotely sensed data have increased significantly over the years. Very often, these data are gathered from different sources and in different formats causing the classification process to be scene-specific. Alternatively, spectral band indices have been developed to emphasize some elements based on spectral characteristics and therefore improving the final classification accuracy. This research applies a multi-disciplinary approach in which genetic programming (GP) and standard unsupervised algorithms are integrated into a single iterative process to develop spectral indices for each element being investigated (such as water, impervious surfaces, dense vegetation, etc). A set of indices formed by mathematical and logical operations of the spectral bands are evolved using genetic operations. The application of non-linear indices enhances the relative spectral difference among the elements investigated improving the clustering capability of the data. The algorithm's ability to generalize provides an alternative to classify multi-temporal data with a single methodology. An example application is given for the water and impervious surface delineation using Landsat MSS, Landsat TM, and Landsat ETM+ imagery. Initial results are comparable to more labor intensive scene-specific supervised classification.

INTRODUCTION

Preprocessing is an additional step introduced in an image analysis schema to minimize the influence of unwanted points and to maximize the influence of desired ones, thus simplifying the computation work needed in later stages of the analysis in order to reach the investigation goals (Paulus and Hornegger, 2003). Preprocessing step are composed of operations such as filtering, enhancing, transformation, or combination of these processes applied to the original measurements to improve the overall schema performance (Sherrah, 1998). Preprocessing is not new in remote sensing and it has been used extensively in the form of morphological filters, principal component analysis, and spectral band indices.

Among the spectral band indices, the vegetation indices are the most developed and used in remote sensing, where more than 20 vegetation indices are in use (Jensen, 2000) (Lillesand and Kiefer, 2000). These vegetation indices apply a pixel-to-pixel operation to create a new value for individual pixels according to some pre-defined function of the spectral values. After the transformation, some features and/or spectral properties become more discernable when compared with the original data. The preprocessing (transformation) function can be expressed mathematically by:

$$I'_{i,j} = T_{i,j}(I_{i,j})$$

where: I - original image, $T()$ - preprocessing function, and I' - preprocessed image

Equation 1

Following the same preprocessing concept of the spectral indices we now propose the development of spectral indices, or preprocessing functions, that use a multi-disciplinary approach involving genetic programming (GP) and standard unsupervised classification algorithms. The use of GP and other evolutionary computation methods for remote sensing applications has increased over the years due to their ability to work with incomplete or missing data, robust classification capabilities, and versatility with different types of data (Daida et al., 1996, Pal et al. 2001, Szymanski et al 2002, and Perkins et al. 2000). Genetic Programming and K-Means algorithms are integrated into a single framework as an optimization tool to develop preprocessing functions for each element being investigated (such as water, impervious surfaces, dense vegetation).

The article presents a brief overview of GP and its main terminology followed by a description of the proposed framework. Initial results from the proposed methods were compared with results from existing supervised classification methods and multi-temporal datasets. Finally the conclusions obtained from these initial experiments are discussed and the future work is described.

PREPROCESSING SCHEMA

Genetic Programming Overview

Genetic programming was inspired by Darwinian theories of biological evolution, which state that individuals that best fit the environment they live in will have a greater chance to survive and therefore pass their genes to the next generation. Conversely, the individuals that are unable to adapt to the environment will likely die out and therefore their genetic material diminishes with time. Over a long period of time a new population of individuals made of superior genetic material is formed.

This evolutionary concept was translated into machine learning as evolutionary computation discipline which is further subdivided into four main approaches: genetic algorithms, genetic programming, evolutionary strategies, and evolutionary programming (Sherrah, 1998). Genetic Programming (GP) is one of the main types of evolutionary computation algorithms and it was first introduced by Koza (1992) to solve problems from different domains. These algorithms search for the optimal hypothesis solution based on measurements provided by the fitness function. These algorithms operate in an iterative mode constantly updating the set of candidate hypotheses (also known as the population) until the most fit hypothesis is found. After each iteration (generation), the set of hypothesis are evaluated based on the fitness function and sorted accordingly. Hypotheses with the highest fitness are then selected to be carried forward to the next new set of hypotheses (new generation). Some of the hypotheses are carried forward with no change (reproduction) while others are susceptible to genetic operations such as crossover, and mutation. The crossover operation produces two new individuals from the parents by copying parts from each parent, while the mutation operation produces small random changes to small parts of the individual (Mitchell, 1997). The simulated evolutionary characteristic of GP is often referred as an optimization algorithm because it uses a combinatorial search approach that is capable of automatically deriving code which proceeds by trial and error repetition (Daida, 1995 and Sherrah, 1998).

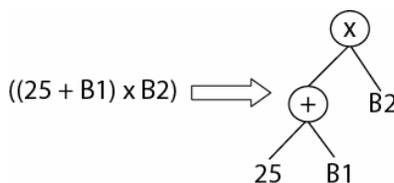


Figure 1. Tree structure representation used in genetic programming.

Genetic programming differs from genetic algorithms and other machine learning algorithms, such as artificial neural network and decision trees, in its hypothesis representation. In GP the individuals are represented by computer programs in tree structures rather than fixed length strings in genetic algorithm or other structure representing set of rules. This alternative representation increases the complexity during the evolutionary process since each individual is represented by a hierarchical tree that constantly changes size and shape. Figure 1 shows an example of a tree structure representation. These individuals are designed to use in the fitness function the results obtained by running the programs one or more times with a set of different inputs (Langdon and Poli, 2002).

Koza (1992) explains that there are five major steps prior to the implementation when designing a genetic programming application to solve a specific problem:

1. Definition of the set of terminals: terminals can be interpreted as the set of arguments and/or inputs used by the functions in the programs constituting the individuals of the population. The items in the terminal set can be constants and variables.
2. Definition of the function set: a set of mathematical functions which will be used as operators in the hypotheses.
3. Definition of the fitness function: the fitness function is the driving force for the evolutionary process. The fitness can be defined many different ways, but is a quantitative measure of the “goodness” of a hypothesis.
4. Definition of the parameters controlling the run: like most computer programs genetic programming algorithms require the definition of some parameters such as: number of population, number of generations, percentage of reproduction, percentage of crossover, percentage of mutation, maximum depth of tree, and others.
5. Definition of the termination criteria: due to the iterative nature of this approach, it is required that a termination criteria to be defined. The establishments of the maximum number of generations and/or the stopping condition when a specific fitness is achieved are the most commons ones.

Evolutionary Framework

Input. Two user provided data sets are used as input into this framework: a subset of the original image and the training points (Figure 2). These training points are used as the reference points (truth points) to which the candidate hypotheses will be compared. In the current application this information is provided by a human analyst in the form of a single band image, referred to as the mask image in this project. The size of the mask image is constrained by computational limitations because it has to be processed numerous times during the evolutionary procedure. The mask image is a subset of the original image and has the same size (number of samples and lines) as the subset of the original image. Each pixel in the mask image can have one of the three values: zero – pixels not considered, one – pixels representing training points where the target element is not found, and two – pixels representing training points where the target element is found.

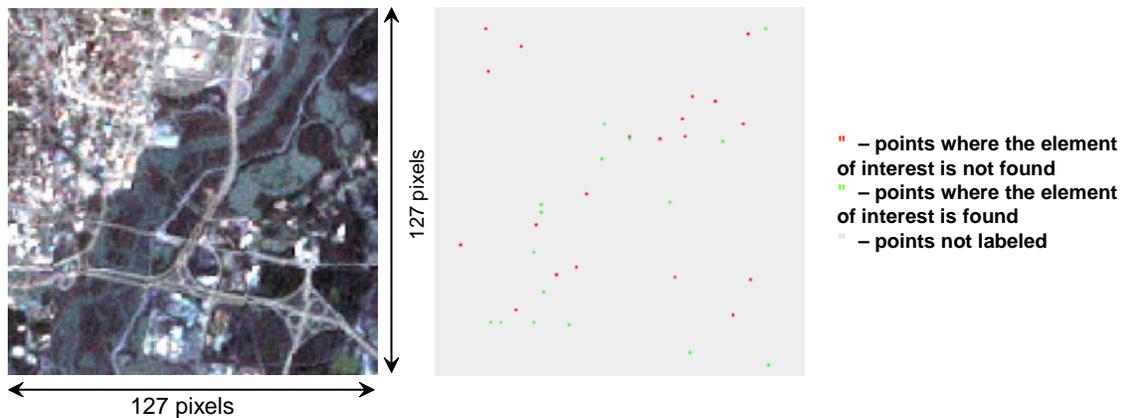


Figure 2. Input datasets: subset of original image (left) and user provided mask image (right) representing the training points used as reference in the evolutionary process.

Fitness Calculation. The fitness assigned to each individual (candidate preprocessing function) is a single number ranging between zero and one, where zero means the least fit score and one the most fit score. This value is calculated by comparing the classified imagery with the mask image expressed mathematically by:

$$fitness = \left(\frac{TP + TN}{TP + TN + FN + FP} \right) \quad \text{Equation 2}$$

where : TP - true positive, TN - true negative, FP - false positive, and FN - false negative

Framework Description. In this hybrid method, the GP algorithm is coupled with a standard unsupervised clustering algorithm to search for the optimal preprocessing function in a learn-from-example schema. Figure 3 shows the flowchart of the overall preprocessing framework. The process starts with the definition of the necessary GP basic parameters, such as: population size P, number of generations G, percentage of crossover PCO, and percentage of mutation PM. The clustering algorithm in Figure 3, will produce a binary imager (presence or absence of the element of interest) of the preprocessed imagery as result of the computation of the discovered candidate preprocessing functions. The K-Means algorithm described by Tou and Gonzalez (1974) was chosen as the clustering algorithm due to its simplicity and low computational cost.

In the first iteration (generation), the GP algorithm randomly generates a set of P candidate preprocessing functions. These candidates preprocessing functions or hypotheses, are represented as tree structures in which two growing rules are adopted regarding the first node and the tree size. The first node (also known as root node) will always be a binary function node to allow the tree to grow in two different directions. The tree size in the first randomly generated population will grow until all the leaves are either constants or variables (terminals).

These candidate preprocessing functions are individually applied to the original image, producing new single banded images (Figure 3 box 5). Each of these preprocessed images is then classified by the K-Means algorithm into a binary classified image (Figure 3 box 7) containing two classes: one representing the class where the target element is not found and another representing the class where the target element is found. Figure 4 shows on the left side an example of the preprocessed imagery and on the right side an example of classified image. This is repeated for each element.

The individual's fitness is calculated by comparing the user defined mask with each individual classified image as described in the previous section. It is expected that all of the individuals in this initial generation will have low fitness values due to the fact that they were randomly generated. However, it is also expected that some individuals will perform better than others, allowing the entire population to be sorted by the fitness values.

After the computation of the fitness values, the algorithm checks the stopping criteria to determine if either of the following conditions has been reached: total number of generations or fitness threshold. If either one of these conditions is achieved, the algorithm outputs the "most fit" individual of the last generation. Conversely, if none of these stopping criteria are met, the process continues by following the principles controlling the biological evolution theory where only the fittest individuals are carried forward for the next generation. In other words, only the hypotheses with the highest fitness values are carried forward to the next generation after genetic operations such as: reproduction, crossover, and mutation are applied. The newly discovered hypotheses are then applied to the original image and the entire process is repeated until the stopping criteria are reached.

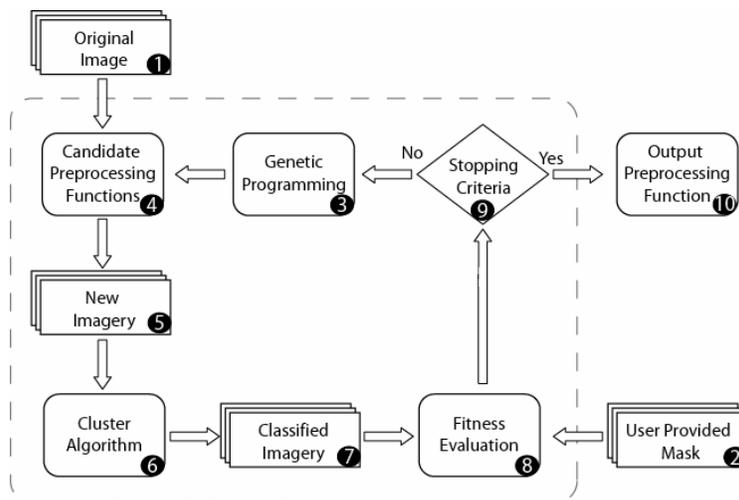


Figure 3. Preprocessing overall framework.

Initial Experiments

In order to check the framework's applicability and generalization capability, the proposed methodology was evaluated by comparing the testing data from different imagery with existing supervised classification methods. Three elements were considered for investigation in this research: water, impervious surfaces, and dense vegetation. It is important to note that these elements were arbitrarily selected and therefore the choice of the number and type of elements varies according to the type of problem being addressed.

Data Preparation. The original imagery used in this project was a Landsat MSS scene of Jackson, Mississippi during the 1983 flood of the Pearl River (USGS, 2005). Figure 2 is a subset of the original image containing 127x127 pixels and six bands. A total of 44 training points were collected by visual inspection of the subset image: 16 for water, 14 for impervious surfaces, and 14 for dense vegetation. As described earlier these points were organized into three different masks: one for each element considered.

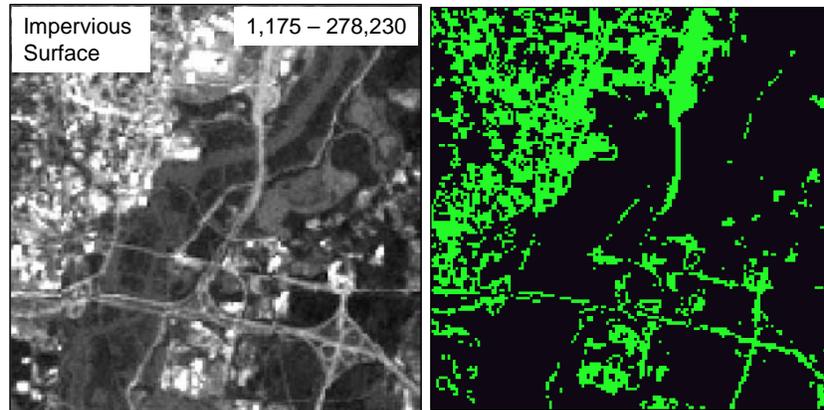


Figure 4. Example of preprocessed image on the left hand side and on the right hand side the same image classified into two classes by with the K-Means algorithm.

In addition to the training data represented by the masks, testing data was also used as the measurement tool to evaluate the overall applicability of the method. Testing points were collected by visual inspection of the 1983 Landsat MSS and 2000 Landsat ETM+ images of Jackson, Mississippi for each element from regions of interest randomly placed over these scenes. This procedure was adopted to spread the collection of testing points throughout the image and therefore to avoid the human bias of only collecting data of a specific element where it is most abundant. The third dataset used for testing purposes was the 1997 Landsat TM of the Louisiana Coastal area which had field data indicating dry land and water.

Table 1. Genetic Programming parameters

Objective:	Find a preprocessing function that will improve the classification by enhancing selected points with similar spectral characteristics and masking the remaining points.
Terminal Set:	Variables: Spectral bands (B1 – B6) Constants: Integer digital numbers (0 – 255)
Function Set:	Binary: (Summation) (Subtraction) (Multiplication) (Safe division) Unary: (Safe logarithm) (Absolute Value) (Safe square root)
Fitness:	Fitness: Comparison of preprocessed imagery with the user provided mask.
Parameters:	Population size: (10 individuals) Number of Generations: (51 – initial plus 50)

Preprocessing Equations Generation. The GP parameters used are shown in Table 01. The terminal set is composed of variables and constants where the variables represent the six possible spectral bands available in the sensor. The thermal band was not considered due to the different spatial resolution. The constant values were limited to the 0-255 range representing the possible band values. The function set is further subdivided into binary and unary functions based on the number of arguments needed by each function. For example, summation is a binary function because it needs two arguments while square root is unary because it only takes only one argument. Note

that division, logarithm and square root have been modified to prevent unwanted situations such as division by zero and complex numbers.

Table 2. Example of preprocessing equations

Water	$((16,722.25) - \sqrt{(BI)}) + (\log(B5) * 142.73)$
Impervious Surfaces	$(\log(BI) * \text{abs}((27.66 * \sqrt{(\log(BI) * B3)})))$
Dense Vegetation	$((\frac{B5}{\log(84.51)}) * (BI - (\text{abs}(B3) - B1)))$

The entire preprocessing step was repeated three times for each of the elements considered (water, impervious surfaces, and dense vegetation) resulting in nine different preprocessing equations. Table 02 shows one example of preprocessing function for each element considered.

PRELIMINARY RESULTS AND DISCUSSION

The developed preprocessing functions were then applied to the original 1983 Landsat scene of Jackson, Mississippi producing nine new single band images (three for each element considered). The spectral values of these images vary considerably among the nine images and within each of them, as shown in the upper part of the left example in Figure 4.

Table 3. Initial accuracy results of preprocessed images classified with K-Means with different number of classes

Preprocessing equations	K-Means number of classes			
	2	4	6	8
Landsat 1983 - Jackson, Mississippi				
Water 01	53.82	73.33	94.86	94.21
Water 02	12.79	80.33	94.21	93.81
Veg 02	32.43	29.42	28.48	28.57
Veg 03	85.18	87.63	83.51	76.03
Imp 01	75.20	87.33	87.33	89.12
Imp 03	77.17	88.68	88.99	88.76
Landsat 2000 - Jackson, Mississippi				
Water 01	88.61	88.61	88.61	87.99
Water 03	77.76	77.76	77.76	88.95

Individual classification of these nine preprocessed images was performed using the K-Means clustering algorithm with different numbers of classes (Table 3). Only two examples of each element are listed due to space limitations. Results have shown that the accuracy (fitness) values vary as the number of classes varies. However, some preprocessing functions resulted in higher accuracy values with higher number of cluster classes while others have peak accuracy values with an intermediate number of classes. Table 3 also shows considerable accuracy differences between the dense vegetation preprocessing functions two and three. The accuracy discrepancy of these two may be due to the randomness involved in the generation of the first population. It is possible that an important part of the solution might not be included in the first generation. In that case, no matter how many generations and/or how many genetic operations performed it is very unlikely that the problem will converge to a solution.

The same 44 training points (16 water, 14 impervious surfaces, and 14 dense vegetation) collected from the 1983 Landsat image and used in the preprocessing function development were also used as seed points for supervised classification algorithms. The Maximum Likelihood and Parallelepiped Maximum Likelihood assumes that the statistics of each class are normally distributed and calculates the probability that each pixel belongs to each class. The parallelepiped method uses a decision rule approach based on boundaries of an n-dimensional

parallelepiped of a multi-spectral image (Richards, 1999). Table 4 shows the accuracy results of these two different clustering algorithms.

The overall accuracy results of the preprocessed imagery are comparable with the values obtained from the supervised classification methods. However, the same preprocessing equations, developed with the 1983 Landsat image, were applied to two different images (1997 – Coastal Louisiana and 2000 – Jackson, MS) and classified with K-Means algorithm with varying number of classes. For the 1997 Landsat scene of the Louisiana Coast only water and non-water field data were available. The accuracy values for the preprocessing equation one applied to this scene range from 92.49 to 93.70 percent with for number of classes between three and five. The results for the 2000 Landsat scene of Jackson, Mississippi are similar to those obtained from the 1983 scene and are also shown in Table 3. The results show the generalization potential of the proposed method for multi-temporal analysis with collection of field data in a single dataset.

Table 4. Landsat 1983 – Jackson, Mississippi accuracy results for supervised classification algorithms using the training points as cluster class seeds

Classification Algorithm	Element	Fitness (%)
Maximum likelihood	impervious surfaces	69.38
	dense vegetation	71.78
	water	92.23
Parallelepiped	impervious surfaces	82.27
	dense vegetation	71.86
	water	93.40

Additionally, experiments were performed to evaluate the possibility of reduction of the correlation between the spectral bands. Table 5A shows the Pearson's correlation values for the original 1983 Landsat data. The high correlation values between spectral bands one, two, three, and six can be interpreted as redundant information. A possible alternative to address this issue is to develop several preprocessing functions and select those that when combined provide the least correlation among them to form a new image. Table 5B shows the correlation values for the nine preprocessing equations developed. In this example the new image would be composed of three bands, one for each element considered. For instance, suppose the goal is the water classification then one possible preprocessing function combination would be “water02”, “veg03”, and “imp02”, since the correlation between “water02” and “veg03” is 0.28 and the correlation between “water02” and “imp02” is 0.26. These three images were stacked into a single image and classified using the same 44 training points as class seeds for the supervised classification procedure resulting in 95.51% accuracy, slight higher than the best result of 94.86%. Figure 5A shows a scatter plot of the stacked preprocessed imagery band 01 (water02) and band 01 (imp02) and Figure 5B shows a thematic image as result of arbitrarily dividing this plot.

Table 5. Spectral bands Pearson's correlation values. (a) 1983 Landsat image. (b) Preprocessed imagery

Correlation	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6
Band 1	1.00					
Band 2	0.91	1.00				
Band 3	0.89	0.95	1.00			
Band 4	0.46	0.58	0.55	1.00		
Band 5	0.64	0.69	0.77	0.72	1.00	
Band 6	0.74	0.76	0.83	0.61	0.95	1.00

(a)

Correlation	water01	water02	water03	imp01	imp02	imp03	veg01	veg02	veg03
water01	1.00								
water02	-0.52	1.00							
water03	0.94	-0.47	1.00						
imp01	-0.56	0.27	-0.68	1.00					
imp02	0.53	-0.26	0.65	-0.99	1.00				
imp03	0.64	-0.31	0.78	-0.93	0.93	1.00			
veg01	0.37	-0.65	0.49	-0.82	0.82	0.76	1.00		
veg02	0.93	-0.47	0.99	-0.76	0.75	0.82	0.57	1.00	
veg03	0.59	-0.28	0.73	-0.91	0.92	0.98	0.75	0.78	1.00

(b)

CONCLUSION

Although the proposed image processing methods will require further work, it is possible to draw some conclusions from the work presented. The discrepancy between the accuracy values of the two different images of dense vegetation imagery suggests a possible weakness of the method. Since the first generation of hypotheses is randomly generated it may be possible that important components of the solution might be left out of the initial generation thus making the entire evolution task more difficult. A possible solution would be to increase the number of individuals in the population to minimize the chances of leaving important parts of the solution out of the initial generation. Alternatively, increasing the percentage of mutation and crossover would also promote diversification at each new generation.

The evolutionary nature of the method produces images with large variation of the spectral values between the images. Figure 5 shows two examples of preprocessed imagery intended to emphasize water created from two different runs using the same input parameters. Both images achieved the accuracy threshold during the evolutionary process however, one has water values as negative values (darker areas on the left side of Figure 5) while the other has the water values as the highest values (lighter areas on the right side of Figure 5) in the scene. For the preprocessing part this may not constitute a problem since the main objective is improve the clustering of water and non-water areas, but if this approach is linked to some automated process (segmentation or feature extraction) this matter has to be addressed.

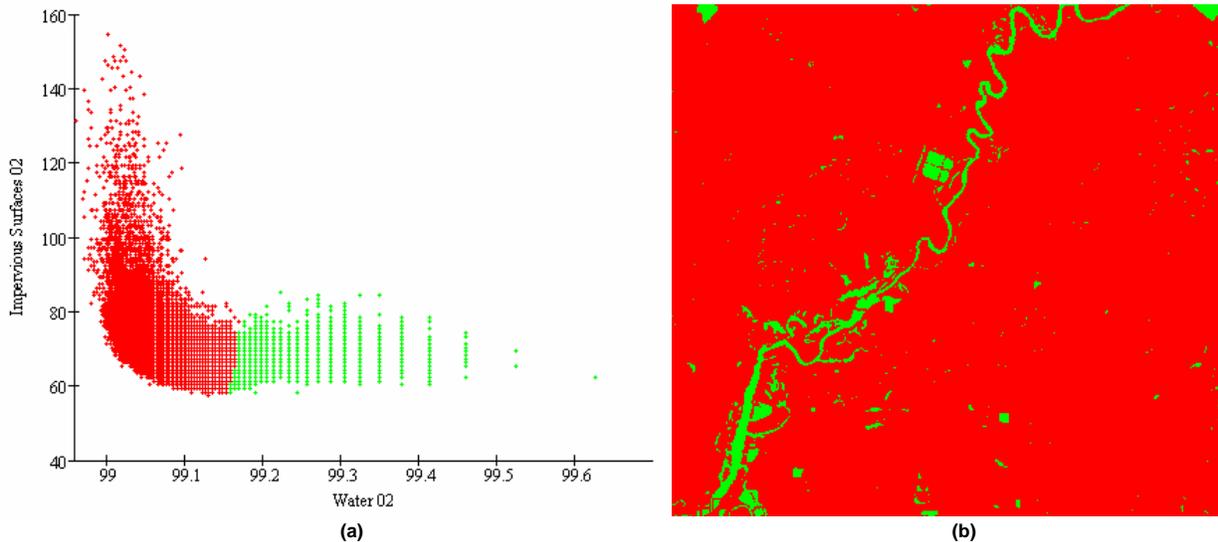


Figure 5. (a) XY Scatter plot between impervious surfaces 2 and water 2 preprocessed imagery. (b) Classified image by arbitrarily dividing the data.

Analysis of the correlation between the preprocessed images revealed the possibility of generating multiple preprocessing functions for the same element, increasing the chances of minimizing the correlation between the elements in the stacked image and therefore improving the final classification results.

The image used to develop the equations represented a unique situation in which the Pearl River had one of the highest flow measured (USGS, 2005) and therefore the flood plains were full with water. But the results from the 1997 Landsat of Louisiana Coastal area and 2000 Landsat of Jackson, Mississippi (with normal flow in the Pearl River, USGS 2005) required the generalization capability of the method. It is important to note that further analysis should be performed with different resolution imagery and different types of elements.

Finally, the framework proposed was not designed to replace or compete with the existing classification methods but rather to be integrated with them to improve the overall result. The main goal was to develop a framework capable of reducing the overall dimensionality of the raw data, highlight user defined spectral characteristics while masking others, and be able to generalize to perform multi-temporal investigations by collecting points in a single dataset. For all of these matters, analysis of the results obtained from the experiments performed, the methodology discussed herein has shown that the method is a step toward these goals.

FUTURE WORK

Sensitivity Analysis

Due to the large number of variables involved in this methodology, it is important to conduct a sensitivity analysis to understand the effect of each variable in the overall process. This sensitivity analysis would include evaluation of the following parameters: GP (variation of the number of generations, the population size, and percentage of mutation and crossover), remote sensing (variation of the spatial and spectral resolution, the number of elements considered, and the size of training data), and function set (variation of the type and number of functions in the function set such as: pixel-to-pixel operations, local convolutions, and global statistics).

Parallel Processing and Morphological Clusters

At the present stage of development the selection of the preprocessed functions to be used in conjunction with others is based on correlation values chosen by the user. This works well with the small number of elements considered, however it poses a problem when there are a large number of elements. A further implementation of the methodology would include a parallel evolutionary process in which all preprocessing functions would be discovered together. In this new approach, it would be possible to take into account the correlation values in the fitness calculation and therefore hybrid fitness would be considered where the most fit individuals would not only have a good match with the mask image but also be the least correlated among the available candidate preprocessing functions.

Additionally improvements would also be achieved using a morphological clustering algorithm rather than algorithm that use the distance to the cluster center. Works have been done with clustering algorithms that use not only the cluster center but also the shape of the cluster and morphological properties of the neighboring pixels (Guha et al. 1998, Jackson 2003, and Karypis, 1999). These techniques may yield improved results in this methodology.

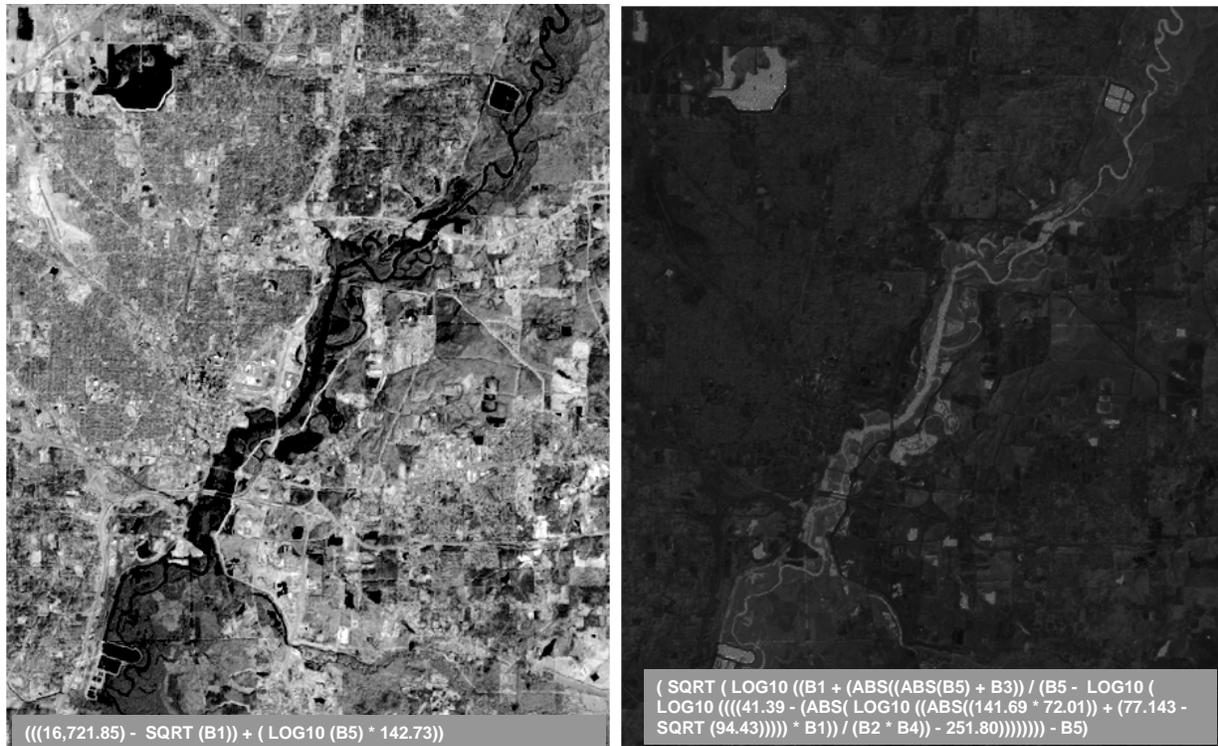


Figure 5. Example of the preprocessed imagery showing the difference in the range of spectral values for two distinct runs using the same input parameters.

ACKNOWLEDGEMENTS

We would like to acknowledge the University of Mississippi Geoinformatics Center (UMGC) for providing the funds and the infrastructure necessary to conduct this research.

REFERENCES

- Daida, J.M., Bersano-Begey, T. F., Ross, J. S., and Vesecky, J. F., (1996). Evolving Feature-Extraction Algorithm: Adapting Genetic Programming for Image Analysis in Geoscience and Remote Sensing, Proceedings of the International Geoscience and Remote Sensing for a Sustainable Future, Washington, USA.
- Daida, J.M., Hommes, J. D., Ross, S. J., and Vesecky, J. F., (1995). Extracting Curvilinear Features from Synthetic Aperture Radar Images of Artic Ice: Algorithm Discovery Using Genetic Programming Paradigm, Proceedings of the 1995 International Geoscience and Remote Sensing Symposium: Quantitative Remote Sensing for Science and Applications, Firenze, Italy.
- Guha S., R. Rastogi, and K. Shim (1998). CURE: An efficient clustering algorithm for large databases. In: Proceedings of ACM SIGMOD International Conference on Management of Data. New York,. ACM. pp. 73 – 84.
- Jackson, P. M, (2002). Building Extraction Using Neuro-Fuzzy Pattern Recognition with Feature Detecting Structures. Master Thesis, The University of Mississippi.
- Jensen, R.J, (2000). *Remote Sensing of the environment: An Earth Resource Perspective*. Prentice Hall, Upper Saddle River, NJ, USA.
- Karypis, G., Han, E. –H., and Kumar, V., (1999). CHAMELEON: A Hierarchical Clustering Algorithm Using Dynamic Modeling. *IEEE Computer*, 32(8), August, pp. 68-75.
- Koza, J.R., (1992). *Genetic Programming - On the Programming of Computers by Means of Natural Selection*, 1. Massachusetts Institute of Technology Press, USA.

- Langdon, W.B. and Poly, R., (2002). *Foundations of Genetic Programming*. Springer-Verlag Berlin, Heidelberg, Germany.
- Lillesand, T. M. and Kiefer, R. W. (2000). *Remote Sensing and Image Interpretation*. Fourth Edition, John Wiley & Sons, Inc, Danvers, MA, USA.
- Mitchell, T.M., (1997). *Machine Learning*. McGraw-Hill International Editions, Singapore.
- Pal, S.K., Bandyopadhyay, S., and Murthy, C. A., (2001). Genetic Classifiers for Remotely Sensed Images: Comparison with Standard Methods. *International Journal of Remote Sensing*, 22(13): 2545-2569.
- Paulus, D.W.R.a.H., J., (2003). *Applied Pattern Recognition - Algorithms and Implementation in C++*. Friedr. Vieweg & Sohn, Germany.
- Perkins, S.J., Theiler, J., Brumby, S. P., Harvey, N. R., Porter, R. B., Szymanski, J. J., and Bloch, J. J., (2000). GENIE: A hybrid genetic algorithm for feature classification in multispectral images, Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation III, San Diego, CA, USA, pp. 52-62.
- Richards, J.A., (1999). *Remote Sensing Digital Image Analysis*. Springer-Verlag, Berlin, pp. 240.
- Sherrah, J., (1998). Automatic Feature Extraction for Pattern Recognition. PhD Thesis, The University of Adelaide, Adelaide, Australia.
- Szymanski, J.P., Brumby, S.P., Pope, P., Eads, D.R., Esch-Mosher, D., Galassi, M., Harvey, N.R., Mc-Culloch, H.D.W., Perkins, S.J., Porter, R., Theiler, J., Young, A.C., Bloch, J.J., and David, N., (2002). Feature extraction from multiple data sources using genetic programming, Proceedings of SPIE -- Volume 4725 Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery VIII, Orlando, FL, USA, pp. 338-345.
- Tou, J.T. and Gonzalez, R. C., (1974). *Pattern Recognition Principles*. Addison-Wesley Publishing Company, Inc., Massachusetts, USA.
- USGS, (2005). United States Geological Survey – Streamflow Stream flow conditions, at: <http://waterdata.usgs.gov/nwis/rt> accessed on January 2005.