Segmentation Analysis of Hyperspectral Data in Green and Road Infrastructures

WIWEKA, Aniati MURNI and M. NATSIR, Indonesia

Key words: Segmentation, Pixel Based, Object Classes

SUMMARY

The land cover consists of natural and manmade objects that effect to the infrastructure environment. In this paper there are two terminologies: combined road and green infrastructure.

The combination of both infrastructure types prevent the environmental damages. They will be catagorized good if their existent give a real contribution.

The monitoring of road and green infrastructures changes associate with their own atribute, form, size and scale of objects. Therefore the one meter resolution hyperspectral data can be used as the alternative monitoring tool.

Because of the pixel based classification constrains always yield salt and paper, we will use spatial based segmentation. By using segmentation, the object classes can be obtained with more precise form and size. The information analysis determine area and volume of object by adding digital elevation model (DEM).

Segmentation Analysis of Hyperspectral Data in Green and Road Infrastructures

WIWEKA, Aniati MURNI and M. NATSIR, Indonesia

1. BACKGROUND

The main objective of the research is interpretate satellite imagery with object extraction such as road and tree as the first duty. For instant there is some complex stuctur in the image consists of many different forms, colors and textures of objects. They will tent to the coarse of efforts to find deeply inside a space of finding object classes. In order to reduce the finding effort from the different information sources must be united. Il will be done by developing a model and fusing together the different sources like image color and cartographic data. In this paper we focus on the fusion of digital elevation model (DEM) with remote sensing imageries. DEM provides information of some special object classes or higher information.

The object classes that are road and vegetation, and other unimportant class like cloud shadows. The shadows cause imagery alter then complicate interpretation processes. While the remote sensing imageries include form, texture and intensity combinations.

The combination of both DEM and the remote sensing data can be utilize for dense areas extractions, and many different classes separation; such as houses, trees, or shadows. The object different in classes needs a model based on interpretation process in order to produce an efficient and stable process.

There are two approach of extracting DEM objects by using elevation information directly or by stereo concept processing. The combination of the high information and the information in the imagery such as low land plateau, houses, and full of vegetation or no vegetation or bar-land is very good by means of high resolution and no mistake in the DEM. By using the simple methods and contour lines, the high object extraction process can be executed, with a land as the reference of elevation.

In the experiment written in the paper, we use CASI and airborne data of The Bogor National Garden (Kebon Raya Bogor), Bogor , West Java, Indonesia.

2. METHODOLOGY

The approach used to resolve the problem discussed in the paper is performed by the methods described in Figure 1 below.

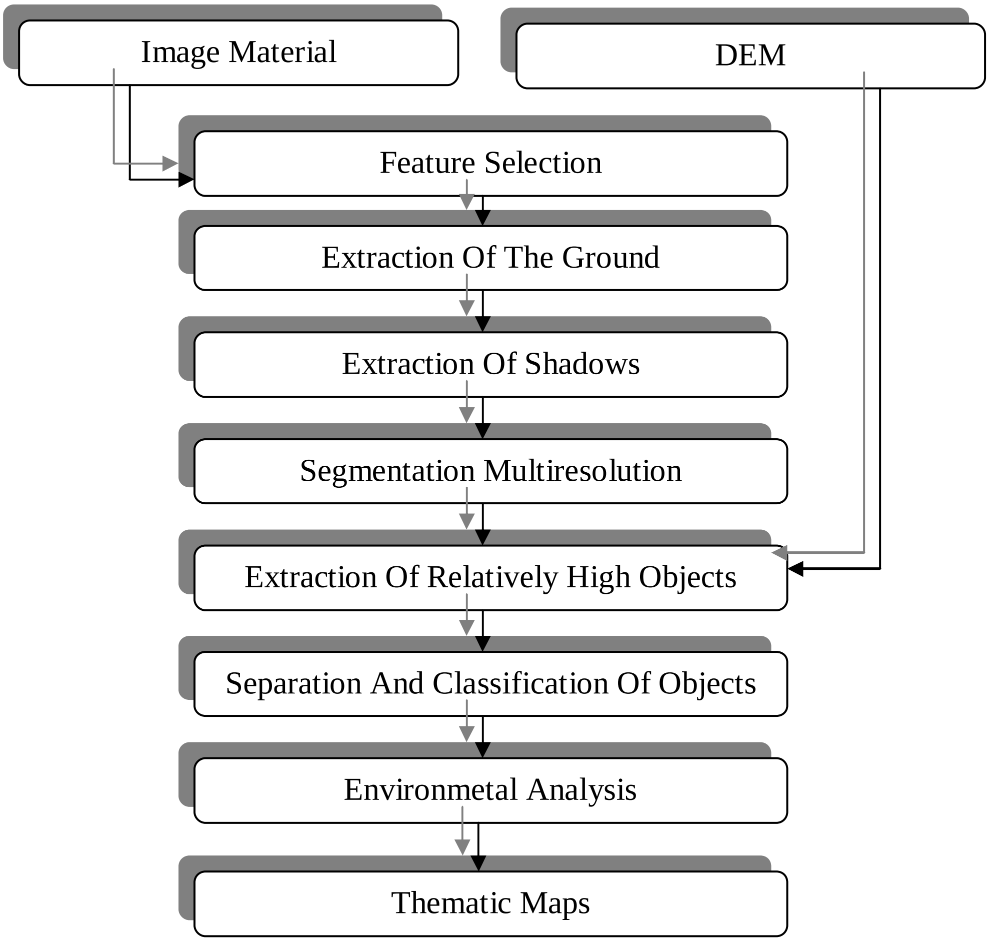


Figure 1: Object Extraction Based on Multiresolution Segmentation

Data used in the experiment written in this paper is the 12 bands CASI data, which has technical specification Dim: 2678 x 1831 x 12 (BIL), Size: (Unsigned Int) 117,682,032 bytes. File Type : ENVI Standard Sensor Type: Unknown Byte Order : Host (Intel) Projection : UTM, Zone 48 South Datum : WGS 84, X Pixel : 1.00000 Meters, Y Pixel : 1.00000 Meters, Wavelength : None Upper Left Corner: 1,1 Description: (Fri Jun 01 21:17:51 2001) . The table 1 bellows describes about the CASI data spectral range used in the experiment.

Geo-correction Methods: Using attitude data (roll, pitch and yaw) recorded in the airplane and data position by GPS differentially corrected. The both data are corrected by bonding every strip one to another (bonding point) and tied to digital topographic map Bakosurtanal of 1:25000 scale, the latest taken photograph (Control point). The final coordinate System expressed in Datum DGN-95 of UTM projection system.

Table 1: CASI Spectral Ranges

445.7nm+/- 5.7nm (rows 266-271) DNSRU:1000.0000000
503.0nm+/- 5.8nm (rows 235-240) DNSRU:1000.0000000
532.9nm+/- 5.8nm (rows 219-224) DNSRU:1000.0000000
571.3nm+/- 4.9nm (rows 199-203) DNSRU:1000.0000000
602.4nm+/- 5.8nm (rows 182-187) DNSRU:1000.0000000
651.6nm+/- 5.8nm (rows 156-161) DNSRU:1000.0000000
674.3nm+/- 5.9nm (rows 144-149) DNSRU:1000.0000000
710.5nm+/- 5.9nm (rows 125-130) DNSRU:1000.0000000
741.0nm+/- 5.9nm (rows 109-114) DNSRU:1000.0000000
775.4nm+/- 5.9nm (rows 91-96) DNSRU:1000.00000000
800.3nm+/- 5.9nm (rows 78-83) DNSRU:1000.00000000
844.3nm+/- 5.9nm (rows 55-60) DNSRU:1000.00000000

The problems that seldom rise in the class identification process in the hyperspectral imageries is the shortage of training data sets. The training data set is desperate as the class identification reference. In the hyperspectral data, the volume size is large that is not balance to the small amount of training data sets. The shortage of training data sets and the abundanse of data volume cause the statistical estimation accuracy go down.

The state where the rise of data dimension that would not be too sure to rise the classification accuracy is called *"curse of dimensionality"* (JAI00). This is then known as *"the peaking phenomenon"*, the larger identity dimension will reduce the classification accuracy. Therefore, although data dimension in the hyperspectral imageries is large, but it does not bee sure that the accuracy is better.

To prevent this phenomena, by reducing the hyper spectral data size so that just includes only the needed information stored and processed. One of the ways to reduce the data dimension is choosing the identity.

The identity selection is one of the important phase in the pattern recoqnation processes. The identity selection basely is input selection that will be used in the pattern recoqnation system (DUD94). The process performed is selection of identity d from given the identity D groups.

The identity selection processes is conducted to prevent the ”*peaking phenomeno*n” that causes the ”*curse of dimensionality”*. While the objective of the identity selection is reducing ”*noise”*, and stay out of irrelevant identities, remove the redundancy, eliminate unused data, and measure the fit in degree between the output of the algorithms and the optimum soulution. (WIW02).

The following is an explanation of identity selection algorithm, that is *Sequential Forward Floating Selection* (SFFS).

The identity selection criterion for measuring identity quality or a group of identity. The identity criterion in this experiment uses range concept usually called *probabilistic distance measure*. The algorithm used in this experiment is *Bhattacharya Distance*, used for measuring the both identity classes spead out.

$$B_D = \frac{1}{4}(\mu_2 - \mu_1)^T [\Sigma_1 + \Sigma_2]^{-1} (\mu_2 - \mu_1) \qquad + \frac{1}{2} \ln \left(\frac{|\Sigma_1 + \Sigma_2|}{2(|\Sigma 1 \| \Sigma 2|)^{\frac{1}{2}}} \right) \quad \ldots\ldots\ldots (1)$$

Where μ_1 and μ_2 are average identity class vectors. Then Σ_1 and Σ_2 are covariant matrices of the identity classes. The identity recoqnation processing result can be seen in the table 2, bellow.

Table 2: Feature Selection Results

File Name	Sub indentity	Band Combination	BD Value	Running time (minute)
bgr_g2	3	1 6 11	$2.4363*10^{-5}$	0.38
	5	1 4 6 10 11	$3.5675*10^{-5}$	0.38
	7	1 3 4 6 9 10 11	$4.6551*10^{-5}$	0.38
	9	1 3 4 6 7 9 10 11 12	$5.09*10^{-5}$	0.38
	11	1 2 3 4 5 6 7 9 10 11 12	$5.4029*10^{-5}$	0.39
bgr_g3	3	3 7 10	$2.6927*10^{-4}$	0.44
	5	2 3 7 8 10	$3 5427*10^{-4}$	0 43
	7	2 3 7 8 10 11 12	$3.9854*10^{-4}$	0.43
	9	2 3 5 7 8 9 10 11 12	$4.2286*10^{-4}$	0.43
	11	1 2 3 5 6 7 8 9 10 11 12	$4.3151*10^{-4}$	0.43

The DEM used in this experiment should has resolution of 2 m, base on the measurement of standard parameters as a realistic data of the automatic interpretation of different imageries. Because of the observation object is land cover, so it can be assume that every object has maximum size of between 20m and 30 m that can utilize the above DEM.

The size is quite large to detect the such a land cover except dent forests or jungles, and very large buildings, and has probability in detecting a flat area even in the mountainous circumtance.

3. SEGMENTATION

According to (DUD04), in the remote sensing techniques, the image segmentation process can be described as :"searching a homogen area in an imagery, then grouping them". There are 3 approach categories): point based (e.g. gray-level thresholding), edge-based (e.g. edge detection techniques) and region-based (e.g. split and merge). In the region based, the imagery objects are reducted with a certain homogenity sizes.

Multiresolution Segmentation, it over a new segmentation called Multiresolution Segmentation (SM). Because SM is such a bottom up technique or region-merging, it is respect to region-base algorithm. SM starts with some considerations of each pixel as a separated object. After considering that every pixel is a separated object, a pair of imagery object are combined to form a larger segment.

The decision in combining the pixels is based on a local homogenity criteria, explained as a similarity between side by side imagery objects. A pair of imagery objects with a smallist improvement in the described combined. The process will be finished when the smallest homogenity improvement passes the userdefined threshold called Scale Parameter (SP). Every bigger value then SP will allow the combination and the consequens the larger object, and vise versa.

The homogenity measurement can be a color combination (spectral values) and structure property (can be devided in smoothness and rigidity). The different Implementation between SP and color or shape combination, can inspire users to create an imagery object net hierarchy.

The segmentation evaluation technique usually can be devide into two categories (supervise and unsupervise). The first category can not be implemented in remote sensing because of an optimum segmentation (field survey segmentation) is difficult to be found. More over, the segmentation evaluation technique has not been tested for overall remote sensing data.

Therefore, for some comparison, it is possible to continue performing classification and then indirectly evaluating segmentation via classification accuracy of the product.

4. ENVIRONMETAL ANALYSIS

As written in the end of the background section, the data used in this experiment are digital CASI and Airborne, and the both data of the Bogor National Garden, are taken in the different times; in 1998 and 2003.

Object classes of this imageries consist of buildings, roads, trees, grass, and pond, based on segmentation cam be measured their areas. According to the title of this report above, this experiment focuses to the changes of roads and trees; since both infrastructures as the main component of good environment and healthy life. Refer to the above function the both object classes must be protected in order that the uses of the both objects can be long.

Table 3: Object Class Area

Number	Object Class	Area	
		CASI	AIRBORNE
1	Buildings	1,410	1,410
2	Roads	10,452	15,908
3	Ponds	5,590	5,590
4	Trees	28,860	59,496
5	Grass	79,564	74,008

From the table 3, we can see that there are changes of length and area roads and a significant of trees canopy area. The ratio between the road infrastructure and the oxygen producer is enough for good environment. Both CASI and Airborne imagery can be seen in Enclosure 1 and Enclosure 2.

5. CONCLUSION

The Multiresolution segmentation capable to provide the best form and size of object class. Therefore it can be recommended that the methods is significant for hyperspectral data segmentation and finaly the accuracy of giving the attribute give a chance to the interpreter ability in finding and recognizing of object classes. The need of increasing this segmentation quality it should be added a field survey, in order provides a more specific class derivative level.

ACKNOWLEGMENT

We would like to thank THEMAP – PT. Comserve Indonesia that allows us to use CASI data in this experiment, PT. Yodya Karya that give us the latest data of experiment location, and BARISTA Geoinformatik that allow us to use e-Cognition

REFERENCES

(DUD04) Duda, O. Richard et al, *Pattern Classification.* New York, John Wiley & Sons, 2001

(JAI00) Jain, A.K. et al, *Statistical Pattern Recognition: A Review*, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, Januari 2000

(WIW02) Wiweka dan Murni Aniati, *Optimasi Feature Selection Data Hiperspektral,* Proposal Disertasi, Universitas Indonesia, 2002

BIOGRAPHICAL NOTES

Name **Aniati Murni**, Born in Magelang of is date of May 29,1948. Education of undergraduate of majors of Technique Electro, Faculty Of Technique, Indonesia University, Education of Master of Grad of Ohio State University USA, year 1981 Education of Doctor of University of Indonesia, Lecturer of Faculty of Computer Science University of Indonesia

Name **M. Natsir**, Born Solo, September20,1951. Education undergraduate Physics Majors grad, Faculty Mathematics And Natural Science University Of Gajah Mada, year 1983. Education S2 Opto Electronic Technique Majors grad, University Of Indonesia, 2000. Researcher, Center Data Processing of Remote Sensing

Name **Wiweka**, Born Kediri 26-1-1964. Education S1 Geodesy Technique Majors grad, Faculty Planning And Civil Engineering, Bandung Institute Of Technology, year 1988. Education S2 Geodesy Technique Majors grad, Bandung Institute Of Technology, 1995. Researcher, Center Application Development and Technological of Remote Sensing

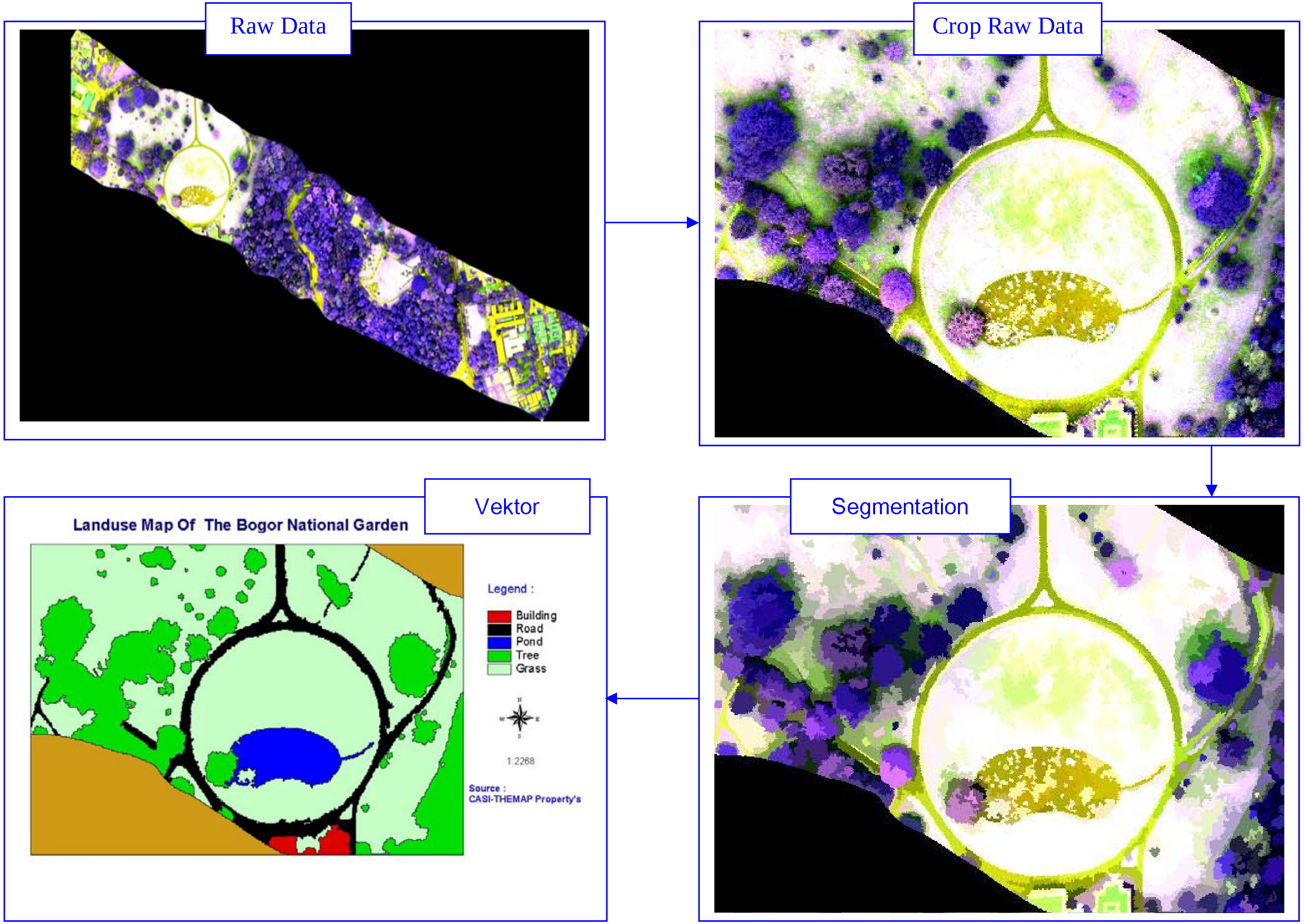
CONTACTS

Aniati Murni
Faculty Of Computer Science
University Of Indonesia
INDONESIA
Tel. + 62 021 786 3419
Fax + 62 021 786 3839
Email: aniati@cs.ui.ac.id

M. Natsir
National Aeronautics and Space Institute
Jl. Lapan 70
Pekayon Pasar Rebo Jakarta Timur 13710
INDONESIA
Tel. + 62 21 871 0786
Fax + 62 21 871 7715
Email: mohnatsir@yahoo.com

Wiweka
National Aeronautics and Space Institute
Jl. Lapan 70
Pekayon Pasar Rebo Jakarta Timur 13710
INDONESIA
Tel. + 62 21 871 0786
Fax + 62 21 871 7715
Email: wiweka@yahoo.com

Enclosure 1
CASI DATA Of THE BOGOR NATIONAL GARDEN



Enclosure 2

AIRBORNE DATA Of THE BOGOR NATIONAL GARDEN

Raw Data

Crop Raw Data

Segmentation

Airborne

Vektor

Landuse Map Of The Bogor National Garden