International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Analysis of Multitemporal Satellite Image Pattern Recognition using Haralick Method to Identify Rice Plant Growth Phase

Chairuddin[a]*, D. Suwardhi[b], L. B. Prasetyo[c], K. Wikantika[d]

[a]Stmik-im, bandung, 40393, Indonesia
[b,d]Geodesy and Geomatics Engineering, Institute of Technology Bandung (ITB), 40373, Indonesia
[c]Forest Resources Conservation Department Forestry Faculty, Bogor Agricultural University, 40776, Indonesia
[a]Email : ch.ruddin@gmail.com
[b]Email : deni@gd.itb.ac.id
[c]Email : lbpras@indo.net.id
[d]Email :ketut@gd.itb.ac.id

Abstract

Pattern recognition is to determine the pattern group or category based on the features possessed by the pattern, the purpose is to distinguish between one object with another object based on the pattern owned by the object in the image. Satellite image is the image data generated from remote sensing technology, satellite image extraction process is needed to interpret the pattern of an object in an image, this is performed so that information from the object being observed can be estimated accurately. Information generated from satellite image data is highly dependent on the image processing that is performed well either by using good model or method to get the results that have a high degree of accuracy.

----

* Corresponding author.

This study focused on the manipulation of multitemporal satellite image data namely Landsat 8 of rice plant growth phase identification to determine pattern group or category based on the features possessed by the image pattern based on the growth phase of rice plant in which the initial process of pattern recognition was through the feature extraction on the image by using gray level co-occurrence matrix (GLCM) and Haralick methods which were then used as inputs of pattern formation. Feature extraction was the first step in determining the parameters as the interpretation of image texture analysis to determine the pattern contained in the image. Based on the analysis results of the comparison of each parameter value in this study, it was found that the results of analysis using angle parameter with the identification of all angles (All) in the feature extraction and the parameter of color intensity, namely Red, Green and Blue (RGB) could be used for object pattern recognition in multitemporal image such as Landsat, the use of these parameters was to obtain more varied information, so that the object pattern information in the image was easier to interpret.

Keywords: Pattern Recognition; GLCM; Haralick; Multitemporal image.

1. Introduction

Remote Sensing technology is a technology transfer concept of image data provider which may be one of the appropriate alternatives in monitoring process, this technology can provide a picture object corresponding to real-world object through image information or picture taken via satellite vehicle.

The main objective of remote sensing is to collect the data of natural resources and environment. This technique usually produces some forms of image further processed and interpreted in order to produce useful data for applications in agriculture, archeology, forestry, geography, geology, planning, and other fields

Utilization of remote sensing technology as the input data in this research is in the field of agriculture, generating the information of rice plant phase pattern in order to monitor by identifying the development process of rice plant. The basic concept is, performing multitemporal satellite image pattern recognition analysis in this case Landsat 8 satellite image using haralick method approach for the identification of rice plant growth phase.

The problem of using satellite image data is at the resolution level generated from the vehicle. Image resolution is the detail level of image, the higher the image resolution, the higher the detail level of the image [7,10].

Image resolution level generated by each satellite is very dependent on the imaging system owned. Although each satellite produces images with different resolution levels, but basically a satellite vehicle producing high-resolution images also has difficulties in interpreting the objects contained therein.

The research conducted by [9,14,17], on image extraction problems can be described as follows:

1. The difficulty of detecting objects in a satellite image because it depends on the level of image resolution in each pixel.
2. Depending on the area of distribution of the object under study, object distribution area must be bigger or

equal to the image capture power.

3. Color or Digital Number and wavelength in an object distribution vary, meaning that in an object distribution under study, it is not always homogeneous .

According to [6,13], the use of remote sensing technology in interpreting clicking food crop objects contains a lot of noise and uncertainty of distribution that cannot be predicted in advance, therefore, the estimation process of food crop object with certain criteria is difficult to predict.

Satellite image extraction process reliability is needed to overcome the problems of object interpretation contained in the satellite image. Therefore, it is needed a reliable model of image processing in order to facilitate the interpretation of object patterns in an image, this is performed so that information from the object being observed can be estimated accurately.

Determination of image patterns performed in this study is to determine the result of information interpretation obtained from system learning process through the pattern input of rice plant object per phase carried out at the stage of pattern recognition, with the feature extraction process in the image to identify object patterns using gray level co-occurrence matrix GLCM and Haralick methods.

Feature extraction is the first step in determining the parameters as image texture analysis interpretation. Texture is a context-dependent feature, that is, texture cannot be defined only from pixels alone but it should be in relation to other pixels in an image region [5]. Texture analysis method used in this study is gray level co-occurrence matrix (GLCM) based on second-order statistics. This co-occurrence matrix was first introduced by [11].

The rationale for using Landsat 8 image in the study is based on the advantages of the satellite image, namely low cost and easily obtained with open data access, having 11 channels, so it is easy to predict the object based on the channel needed, gray level (digital number) having long interval that is: 0-4096 with the sensitivity of each pixel having quantification of 12 bit, so that it will be easier to differentiate objects display on earth surface, thereby reducing interpretation errors, as well as having a scan area of 170 km x 183 km, Time series [19].

The weakness of Landsat image use in this study is the degree of spatial resolution of 30x30 m and temporal resolution of 16 days [19], it is because the plant area is varied and the planting time is different. Solutions used to overcome these problems are:

a. The object under study is the distribution of rice field that has rice plant prevalence rates of more than 30x30 m in a study area, the spatial resolution of 30x30 meter per-pixel does not become a problem.
b. Temporal resolution of Landsat 8 satellite is 16 days, it is very possible considering the monitoring of each phase of the rice plant distribution growth is 1 month.

The concept of satellite image pattern recognition analysis model in this research aims to generate reliable texture analysis value that can be used as input parameters for the process of multitemporal image extraction[16]. It is expected that the analysis model concept can provide a good parameter to generate the information of multitemporal satellite image extraction value namely Landsat 8 and can be the optimal solution in order to carry out monitoring to identify rice plant object per growth phase, as well as to minimize the geographical factor constraint and limitated facilities so that the costs incurred can be minimize.

2. Methods

2.1. Research stages

The research stages were:

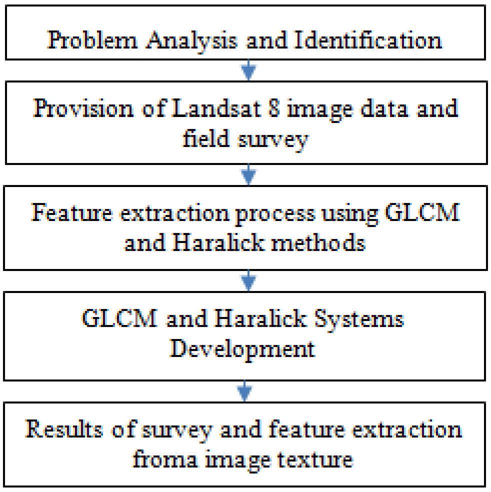


Figure 1: Research stages

2.2. Problem analysis and identification

Based on the research objective, that was to produce texture analysis value of satellite image data samples per phase growth of rice plants that can be used as input parameters for the extraction process, the problem analysis and identification conducted at this stage was aimed to determine the scope of the study of texture analysis for pattern recognition, as well as problem identification from the scope of the study [3,15,16].

2.3. Provision of image data

The stage of image data provision was carried out as image extraction system learning initial input to recognize the form of rice plant per-phase derived from the brightness of the color produced from satellite images sample[1,2]. The satellite images sample obtained was the result of cropping based on the coordinates from the results of field survey, the data would be reference data used as training and testing data used as inputs by the extraction system in order to generate information about the distribution of rice plants per-phase[10,12].

The research area as the sample of image and GPS coordinate was Gedebage, West Java, Bandung, Indonesia. The following figure shows the Landsat image data and the coordinates of the survey result.

Reference data of rice plants per growth phase used as training and testing data for the system, was taken from the determination of the research area with the area of rice plant distribution per phase, in which in this study the image used was Landsat satellite image with the spatial resolution of 30m x 30m. To make the field data taken was in accordance with the criteria of Landsat 8 image, the field data taken had to be the area greater than or equal to 22,500 m2.

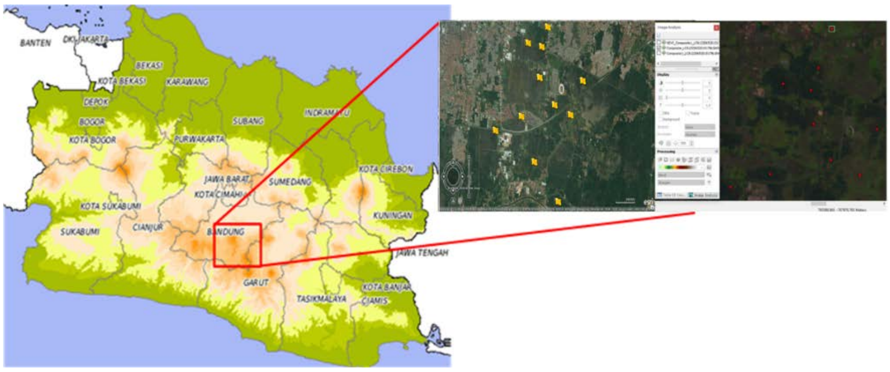


Figure 2: Research area

The stages of reference image data collection process are:

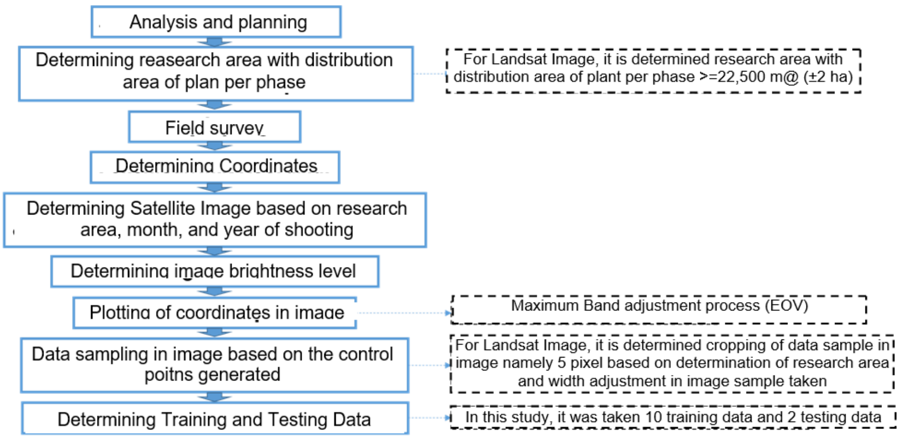


Figure 3: Stages of image data provision

1. Analysis and planning, the initial stages of analysis and planning were the process of satellite image data provision used as reference data for the system identification and learning process. The next was the needs identification of hardware, software, and site survey.
2. Determining the research area with the distribution of plant per-phase. Based on the identification results of Landsat image data needs with a spatial resolution of 30m x 30m, it was determined that the area of rice plant object per growth phase to be identified was 22,500m^2 or five pixels if converted into Landsat image resolution.

The determination was based on the results of analysis and initial survey on the pattern of early planting (seeding) of different rice plants object, in which the average difference of early planting (seeding) was three weeks.

3. Field Survey, survey location determination was based on the analysis result of rice plant object distribution area determination, that was bigger than or equal to 22,500m^2. Based on this, the focus of survey location determination was the distribution rice plant object per phase with an area bigger than or equal to 22,500m^2 with the difference in the pattern of early planting (seeding) in the location area one to three weeks.
4. Determining the coordinates was carried out at the central point of the area to be used as a reference point of data collection in the satellite image; this was carried out so that the data collection of five pixels in the image referred to a predetermined reference point during the field survey.
5. Determining satellite image data was used as training data and testing data based on the research area, months and years of shooting when the field survey, so that the data used as input reference data from the system was in accordance with the growth phase of rice plants according to the survey results .
6. Determining the level of image brightness, Landsat 8 image sharpness increase was carried out by pan-sharpening, i.e. by combining band 8 (panchromatic) into RGB composite image. Band 8 had significance in image processing because of its related advantages of spatial resolution. Of the 11 bands owned by Landsat 8, band 8 spatial resolution was the highest, at 15 m per pixel. Compare with Band 1 to 6, which only had a spatial resolution of 30 m [19].
7. Plotting of the coordinates in the image was carried out to determine the reference point of data collection in the image used as input data i.e. training data and testing data.
8. Sample data collection in the image for training and testing data was determined by five pixel based on the results of reference point plotting on the satellite image. The determination of five pixels was based on the coverage of research area , so that the information of each pixel of the image can be utilized as the material of identifying the object under study.
9. Determination of training and testing data was carried out by taking 10 image samples for training data and two image samples for training data, the data was obtained from the cropping of five pixels of each phase of different plants growth and from the time of the different image collection (shooting).

2.4. The process of feature extraction

2.4.1. Pattern identification process model with texture analysis

Pattern identification model and texture analysis were designed to determine the stages of development of GLCM application systems using six features of Haralick texture identification method. The components used in the development of the system model were:

Input: Yield Data, Data cropping of satellite image of each Plant Growth Phase at a certain coordinate point

Output:

1. Texture identification from 6 features of Haralick method for each phase of growth

2. Parameter for Training Data and Testing Data

Objectives:

1. To identify the difference in texture of Rice Growth in each phase of growth
2. As the input or input data for pattern recognition

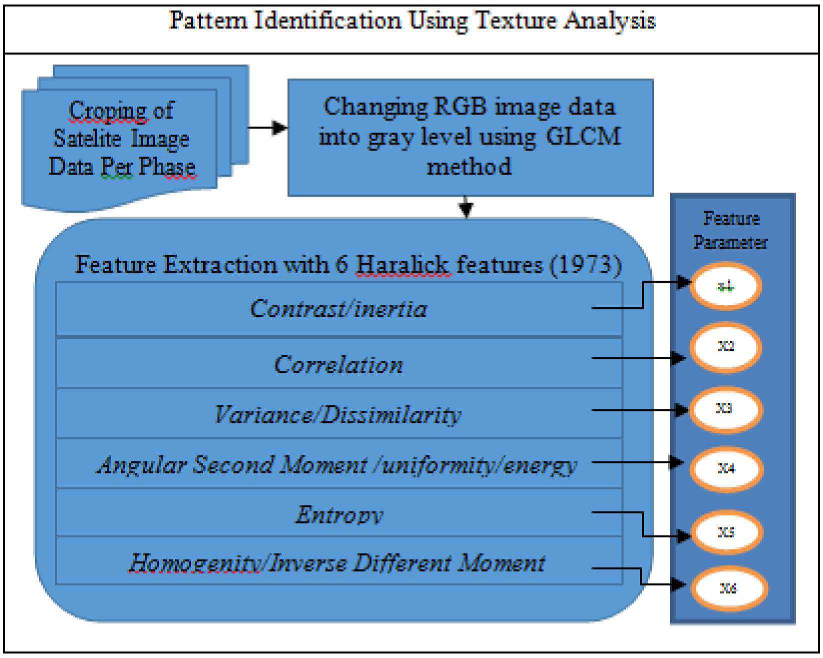


Figure 4: Texture analysis system model

2.4.2. Feature extraction

Feature extraction was the first step in determining parameters as the interpretation of image texture analysis. Texture was a context-dependent feature, which was, texture could not be defined only from pixels alone but it had to be in relation to other pixels in an image region [5].

Texture analysis method used in this study was gray level co-occurrence matrix (GLCM) based on second-order statistics. Co-occurrence matrix was first introduced by [11].

According to [11], co-occurrence means joint events, i.e. the occurrence number of one level of pixel value neighboring with one another level pixel values within (*d*) and specific orientation angle (θ). Distance was expressed in pixels and orientation was expressed in degrees. Orientation was formed in four directions of angle with the intervals of 45°, 0°, namely 45°, 90°, and 135°, while the distance between pixels was usually set at 1 pixel.

This process related to the quantization of image characteristic into a set of corresponding features values representing the neighborhood relationship between pixels in various orientation direction and spatial distance to extract features used as image analysis resulted from remote sensing.

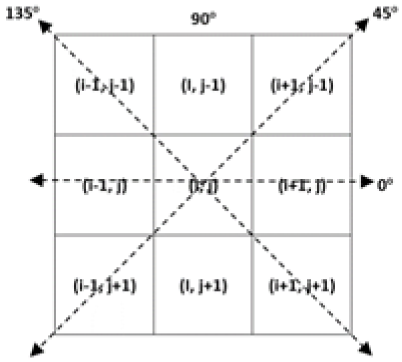


Figure 5: Co-occurrence matrix

Haralick method proposed 14 sizes (features), but Connors and Harlow in 1980s examined that of the 14 features proposed by Haralick, only five of which were usually used. These five features were: energy, entropy, correlation, homogeneity, and inertia [4,8]. In this study, the six features used were among others:

a. Contrast / inertia

Shows the size of deployment (moment of inertia) of image matrix elements. If located far from the main diagonal, the contrast value is big. Visually, the contrast value is variation measure among gray degrees of an image area.

$$CON = \sum_{k} k^2 \left[\sum_{i} \sum_{j} p(i,j) \right]_{|i-j|=k} \quad (1)$$

b. Correlation

Shows the size of linear dependence of image gray degree so as to provide guidance for linear structures in the image.

$$COR = \frac{\sum_{i} \sum_{j} (ij).p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (2)$$

c. Variance / dissimilarity

Shows variations of concurrence matrix elements. Image with a small degree of gray transition will have little variations as well.

$$VAR = \sum_{i} \sum_{j} (i - \mu_x)(j - \mu_y) p(i,j) \quad (3)$$

d. Angular Second Moment / uniformity / energy

Shows the size of image homogeneity property.

$$ASM = \sum_i \sum_j \{p(i,j)\}^2 \qquad (4)$$

where *p (i, j)* states value in row i and column j in the concurrence matrix.

e. Entropy

Shows the size of form irregularity. ENT big value is for the image with even gray degrees transition and of ENT small value if the structure of the image is irregular (varied).

$$ENT_2 = -\sum_i \sum_j p(i,j) \cdot {}^2\log p(i,j) \qquad (5)$$

f. Homogeneity / Inverse Different Moment

Shows the homogeneity of image with similar gray degree. Homogeneous image will have bigger IDM value.

$$IDM = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j) \qquad (6)$$

2.5. System development

GLCM application system was built using Matlab, the process of system development followed the design of texture analysis system model, with the stages of system establishing process from satellite image data input resulted from five pixels cropping, Gray level determination from the image formation, concurrence value determination that meant joint events, namely the occurrences number determination of one level of pixel value neighboring with another one pixel value level within the distance (d) and specific angle orientation (θ). The distance was expressed in pixels and orientation was expressed in degrees, then the result of distance determination was processed to extract the image data in the form of a matrix co-occurrence used as the image analysis in Haralick method features [18].

The features used in this study are: Contrast, Correlation, Dissimilarity, Energy, Entropy, Homogeneous. The features used were customized based on the research needs.

The system also detected the color parameter or value contained in the image based on the basic color intensity level, namely: Red, Green and Blue (RGB). The color intensity was represented in two forms, namely parameter and histogram (frequency and range) values.

Here is a figure of system menu built

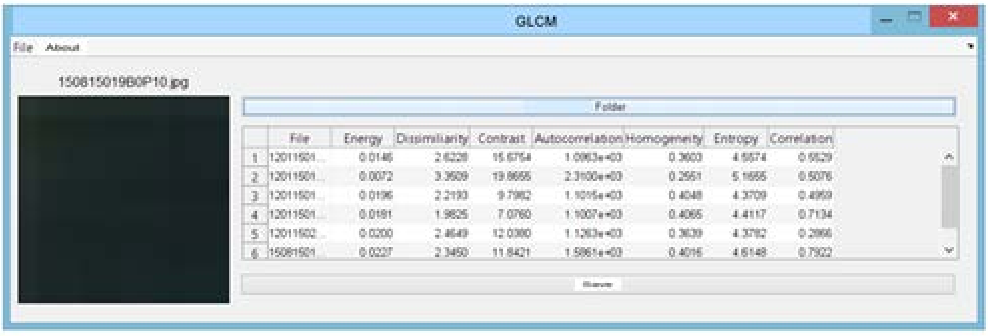


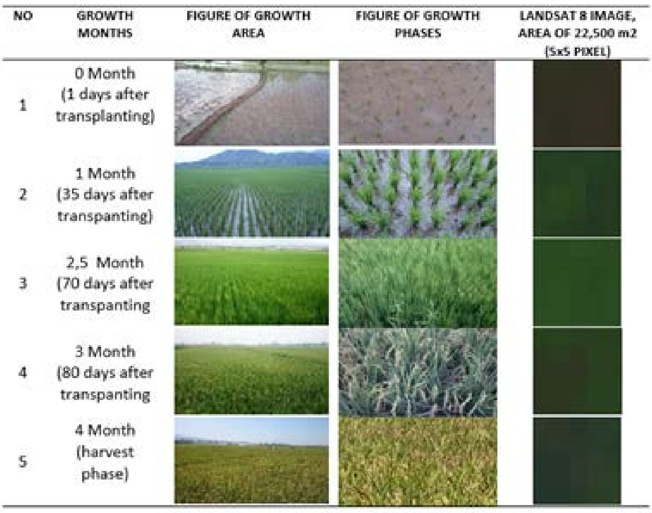
Figure 6: GLCM System Menu

2.6. The results of survey and image extraction

Based on the observation made in the field, it produced images of plant growth phase and a control point from GPS. GPS control point was plotted on Landsat 8 image obtained based on the adjustment of the recording date. 5x5 pixel reference image data assumed by 22,500M^2 area, was obtained from the cropping based on the coordinates and growth phase.

The data presented was a sampling of some generated image, taken at random by the growth phase. The following table shows the growth phase image and reference image based on the results of a survey conducted

Tabel 1: Field survey data and image cropping



NO	GROWTH MONTHS	FIGURE OF GROWTH AREA	FIGURE OF GROWTH PHASES	LANDSAT 8 IMAGE, AREA OF 22,500 m2 (5x5 PIXEL)
1	0 Month (1 days after transplanting)			
2	1 Month (35 days after transpanting)			
3	2,5 Month (70 days after transpanting			
4	3 Month (80 days after transpanting			
5	4 Month (harvest phase)			

The cropping of Landsat 8 image was based on the growth phase used as the reference image,

The cropping reference image of five times five pixels taken based on the growth phase was used as input for the process of GLCM and feature extraction of Haralick method performed by the system. System output generated parameter values of the reference image analysis.

Tabel 2: System output parameter values

File map	Fase	Degree	GLCM						RGB					
									Red		Green		Blue	
			Contrast	Correlation	Dissimilarity	Energy	Entropy	Homogen	MeanR	StdR	MeanG	StdG	MeanB	StdB
		0	176,21	0,51	0,81	0,18	1,91	0,99						
		45	265,18	0,03	1,21	0,18	1,95	0,98						
120115011	0	90	88,81	0,80	0,41	0,19	1,87	0,99	25,34	16,65	34,63	16,11	31,79	16,28
		135	265,17	0,03	1,21	0,18	1,95	0,98						
		ALL	198,78	0,42	0,91	0,18	1,92	0,99						
		0	80,88	0,68	0,41	0,24	1,68	0,99						
		45	161,83	0,03	0,81	0,24	1,76	0,98						
040215011	1	90	80,95	0,67	0,41	0,24	1,69	0,99	29,72	13,24	48,32	12,47	38,83	12,91
		135	161,97	0,03	0,81	0,24	1,76	0,98						
		ALL	121,37	0,42	0,61	0,24	1,73	0,98						
		0	67,34	0,67	0,37	0,27	1,52	0,99						
		45	134,66	0,02	0,73	0,27	1,56	0,98						
270515015	2	90	67,34	0,67	0,36	0,28	1,49	0,99	42,21	12,09	71,04	11,07	27,59	12,61
		135	134,80	0,02	0,73	0,27	1,56	0,98						
		ALL	101,01	0,41	0,55	0,27	1,54	0,98						
		0	74,30	0,66	0,39	0,25	1,60	0,99						
		45	145,73	0,02	0,76	0,24	1,66	0,98						
280615015	3	90	71,45	0,68	0,38	0,25	1,59	0,99	43,05	12,50	59,87	12,54	32,14	12,73
		135	145,87	0,02	0,76	0,24	1,66	0,98						
		ALL	109,31	0,41	0,57	0,24	1,63	0,98						
		0	74,96	0,68	0,40	0,12	2,34	0,98						
		45	147,31	0,10	0,79	0,12	2,42	0,97						
040215015	4	90	72,36	0,69	0,39	0,12	2,32	0,99	41,66	12,85	57,18	12,33	41,51	12,49
		135	147,45	0,10	0,79	0,12	2,42	0,97						
		ALL	110,49	0,45	0,59	0,12	2,38	0,98						

3. Results and discussion

Based on the value of each parameter generated by the system, then it is performed standardization process by finding the mean of each parameter of extraction feature and color intensity feature that is RGB.

The results obtained from the value standardization process are, as shown in the following table,

Table 3: Results of standardization process

Month	Mean_0	Mean_45	Mean_90	Mean_135	Mean_All	Mean_RGB
Month 0	30.10	44.92	15.51	44.92	33.87	30.59
Month 1	14.15	27.61	14.16	27.63	20.89	38.96
Month 2	11.86	23.04	11.85	23.06	17.46	46.95
Month 3	13.03	24.90	12.56	24.92	18.86	45.02
Month 4	13.25	25.28	12.81	25.31	19.17	46.78

From the table, it can be seen the fluctuating value generated from every angle parameter based on the degree of 0, 45, 90, 135 and all angles (All). It appears that the use of the degree with the reading and determination of all the angle and direction (All), is more stable compared to only using one of the four corners, that were 0, 45, 90 and 135.

The use of parameter values of color intensity, that is RGB, is expected to add the process of identifying the object information contained in the image.

The following is the graph showing fluctuating extraction parameter value generated.

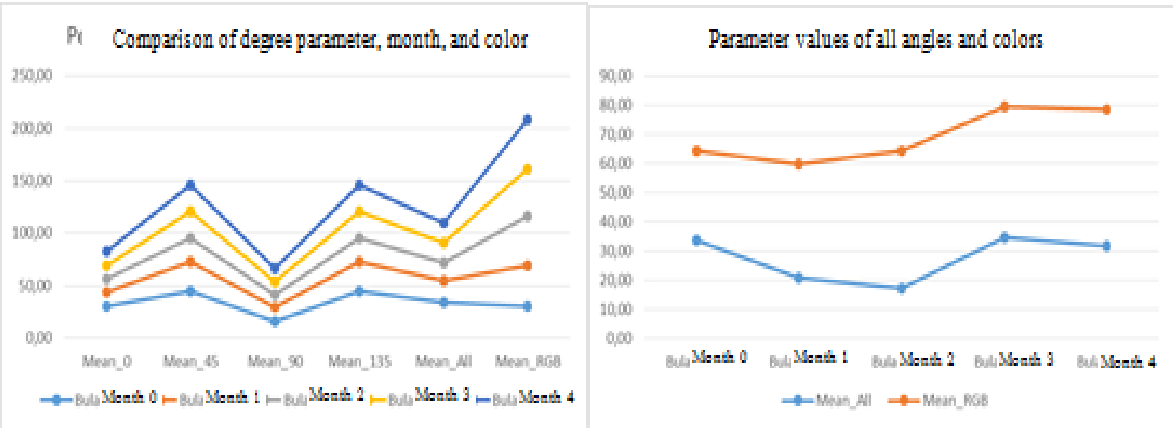


Figure 7: Comparison of the extraction parameter value

a) Comparison of the parameter values per phase of growth based on the angle parameter and RGB

b) Comparison of the angle parameter that is all angles (all) and RGB based on the growth phase

The resulting graph shows the comparison value of each extraction parameter, so that the results of the analysis of each process can be identified properly. In the graph (a), it is showed that the angle parameter by reading all the angles represented by the value of Mean_All looks more balanced and more stable compared to the value of another angle. The balance of the mean of all angles (Mean_All) compared to the RGB color intensity parameter on the graph (b), looks aligned and has the same pattern. If both of these parameters are used for a process, they are expected to be mutually reinforcing and mutually supportive.

4. Conclusion

Based on the analysis results of comparison of each parameter value, it is obtained that the analysis results can be expressed:

1. Parameter of angle or degree with the reading and determination of all the angle and direction (All) can be used for the identification of objects in multitemporal image such as Landsat, compared to only using one of the four angles of 0, 45, 90 and 135, because the reading for all angles at each pixel will get more varied information, so that the information of object pattern in the image is easier to perform

the comparison process.

2. Parameter of color intensity that is Red, Green and Blue (RGB) is best used as a complement and reinforcement value in the process of identification of objects in the image, because the pattern of the object in the image can be more recognizable than the mean of intensity of the color resulted from the object in each pixel,
3. The use of GLCM and Haralick method can be used for pattern recognition analysis of multitemporal satellite image for the identification of rice growth phase with an area 22,500m^2 on the ground and five times five pixels on Landsat 8 satellite image.

References

[1] D.Haboudane, J.R. Miller, E.Pattey, P.J.Zarco-Tejada, I.B. Starchan, “Hyperspectral Vegetation Indices dan Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in theContext of Precision Agriculture”, Remote Sensing of Environment, Vol. 90, hal. **337-352**, 2004.

[2] D.Stathakis, I.Kanellopoulos,”Global Elevation Ancillary Data for Land-use Classification Using Granular Neural Networks”, Photogrammetric Engineering & Remote Sensing, Vol. 74, No. 1, January 2008, American Society for Photogrammetry and Remote Sensing

[3] D.Stroppiana,M.Boschetti,P.A.Brivio,S.Bocchi,“Remotely Sensed Estimation of Rice Nitrogen Concentration for Forcing Crop Growth Models”, Italian Journal of Agrometeorology, Vol. 3, hal. 50-57. 2006.

[4] E.Kulak, “Analysis of Textural Image Features for Content Based Retrieval”, Thesis, Sabanci University,2002.

[5] K.Hauta., Markku, “Computational Techniques for Spectral Image Analysis”, Thesis, Lappeenranta University of Technology Finland,1999.

[6] N.Widyastuti,A.Hamzah,”Penggunaan Algoritma Genetika Dalam Peningkatan Kinerja Fuzzy Clustering Untuk Pengenalan Pola”, Berkala MIPA, 17(2), Mei 2007, Fakultas Teknologi Industri, Institut Sains & Teknologi AKPRIND Yogyakarta.

[7] P.Mahinda, Pathegama., G.Özdemir.”Edge-end Pixel Extraction for Edge-based Image Segmentation”, World Academy of Science, Engineering and Technology, 2005.

[8] R. Conners, C. Harlow, “Toward a structural textural analyser based on statistical methods”. Computer Graphics and Image Processing 12 (1980) 224-256.

[9] R.C. Gonzales, R.E.Woods,. “Digital Image Processing” Third Edition, Prentice Hall, 2008.

[10] R.G.Congalton, K.Green.”Assessing The Accuracy of Remotely Sensed Data”: Principles and Practices 2nd Ed. CRC Press, Taylor and Francis Group New York, 1999.

[11] RM.Haralick, K.Shanmugam, I.H.Dinstein.”Textural features for image classification”. IEEE Transac, SMC-3:610-621,1973.

[12] S.K.Bandyopadhyay, P.Sanyal, “Application of Intelligent Techniques towards Improvement of Crop Productivity”, International Journal of Engineering Science and Technology (IJEST), 2011.

[13] S.Mulyono,E.Piantari, M.I.Fanany,T.Basaruddin.”Pemilihan Fitur Citra Hiperspektral Hymap Dan Model Prediksi Panen Padi Menggunakan Algoritma Genetika Dan Regresi Komponen Utama”, Badan Pengkajian dan Penerapan Teknologi(BPPT), 2011.

[14] S.Ying,H.Guo-Jin,"Segmentation of High-resolution Remote Sensing Image Based on Marker-based Watershed Algorithm". IEEE, Fifth International Conference on Fuzzy Systems and Knowledge Discovery, 271-276. 2008.

[15] T.Blaschke,"Object based image analysis for remote sensing", ISPRS Journal of Photogrammetry and Remote Sensing 65 (2010) **2_16**

[16] Y. Bedard,"Uncertainties in Digital Image Processing". Proceeding Auto Carto VIIK p.**175**, 1987.

[17] Z.Yun,F.Xuezhi,L.Xinghua,. Segmentation on multispectral remote sensing image using watershed transformation. IEEE, Congress on Image and Signal Processing, 773-777,2008.

Internet

[18] M.Syarif, GLCM, http://www.softscients.web.id/, [Des.3, 2015]

[19] NASA(2015),Landsat8,http://landsat.gsfc.nasa.gov/news/news_archive/news_o429.html,[Agu.24,2015]