

LABELING OF CLUSTERS BASED ON CRITICAL ANALYSIS OF TEXTURE MEASURES

D. Singh ^{a,*} and V. Vijaya Chamundeeswari ^b

^{a,*}Department of Electronics & Communication Engineering, Indian Institute of Technology Roorkee, Roorkee, India

^bDepartment of Computer science Engineering, Velammal Engineering College, Chennai, India

KEY WORDS: Clusters, Backscattering Coefficient, Surface Roughness, Labeling, SAR.

ABSTRACT:

It is well known that unsupervised classification of a single polarized SAR image is accomplished mainly by two steps i.e., (i) Clustering the SAR image into groups or clusters on the basis of backscattering coefficient and textures present in SAR image, and (ii) Labeling the various clusters in their respective class (For example, land cover types such as water, urban, agriculture or any other areas). In this context, labeling is termed as naming the various clusters or groups of pixels according to nature of the terrain as a certain land cover type it belongs to. Labeling of various clusters is a crucial and important aspect to identify various clusters in their original class (here, land cover class is assumed as class, whereas, in general, class may refer to any group of targets). It is still a challenging task to label the cluster without any a priori information. So, it is important to develop such a technique by which clusters can be labeled according to their class. Therefore, focus in this paper is to induct surface roughness with backscattering coefficient to label various clusters (i.e. major land cover types). We have proposed an empirical relation to estimate roughness parameters from the SAR image. Labeling of clusters was carried out on the basis of roughness parameters and backscattering coefficient.

I INSTRUCTIONS

SAR image is widely recognized for its potential in investigating soil and vegetation properties (Anitha et al., 2006; Prakash and Singh, 2008). Land cover classification has been attempted by Dobson (Dobson et al., 1995) by estimating terrain attributes by analyzing the sensitivity of backscatter coefficient to various parameters ranging from surface roughness, canopy architecture, soil moisture content, and aboveground biomass of vegetation with its moisture status. Although electromagnetic scattering from surfaces have been extensively studied for a long time, retrieval of surface parameters could be done for specific conditions like that of bare soil, or agriculture areas (Qiang et al., 2007). For studying the separability of urban regions in SAR image, radar response from various urban features is to be studied. It has been shown that some typical urban features like pavement, grass, and buildings can be separated on the basis of their average vertical dimensions (Xia and Henderson, 1997).

There is a need to include some measures from SAR image to label the clusters. Roughness is one of the important parameter to describe the nature of the terrain or surface and it is defined as vertical deviations of a surface from a reference level and it may be helpful to label the various clusters into land cover types. Therefore, labeling the clusters can be attempted using the surface roughness parameters. In the case of urban areas where city streets or buildings are lined up in such a way that the incoming radar pulses are able to bounce off the streets and bounce again off the buildings (called a double bounce), then directly back towards the radar, they appear very bright in SAR images. Agriculture areas have intermediate backscattering coefficient. The surface fluctuation of agriculture areas is also intermediate to flat surfaces and urban areas. Hence, use of backscattering coefficient along with surface roughness may help in labeling various clusters.

II Developing Surface Roughness from texture measures

2.1 Generation of synthetic images with varying roughness measures

The main aspect of characterizing the roughness of a surface is the choice of suitable parameters to represent roughness. For SAR images, vertical size of the scatterers and the distribution of these scatterers in the horizontal plane are important for studying the surface roughness. There are two statistical parameters known as the standard deviation of surface height variation, termed as root mean square (RMS) height and abbreviated as 's' and the correlation length, denoted by 'l'. RMS height, standard deviation of surface height, represented by 's' is a measure of vertical roughness whereas correlation length, representing correlation on horizontal plane, denoted by 'l', is a measure of horizontal roughness. 's' and 'l' are taken in units of cm.

The surface profile generation algorithm by Fung and Chen (1985) is modified for 2D surface profile.

The surface profile height Z is described by

$$Z(k) = \sum_{j=-M}^M W(j)X(j+k) \quad (1)$$

where Z(k) is the surface height distribution, X(i) is a Gaussian random deviate with zero mean and unit variance, and W(j) is the weight function. The surface is characterized by Gaussian correlation function for which the weight function is defined as (Pant et al, 2010)

$$W(j) = s \left(\frac{2\Delta x}{l\sqrt{\pi}} \right)^{1/2} \exp \left[-2 \left(\frac{j\Delta x}{l} \right)^2 \right] \quad (2)$$

Δx is the sampling distance, and it is considered as unity since discrete profiles and consequently discrete image is considered. Various synthetic images have been generated by varying RMS height, 's' from 0.1 to 5.0 in steps of 0.1 and correlation length 'l' from 0.5 to 15 in steps of 0.5. This gives rise to 50 values of 's' and 30 values of 'l'. For each value of 's', 30 values of 'l' are used to generate 30 synthetic images. Hence for 50 different values of 's', (30×50) i.e., 1500 synthetic images are generated

from all combinations of 's' and 'l'. These generated synthetic images are used for further analysis.

2.2. Analysis of texture measures for surfaces with varying roughness parameters

Texture measures namely mean, variance, semivariogram, lacunarity, weighted rank fill ratio and wavelet components are computed for various generated synthetic images (approximately 1500 images) for different 's' and 'l'. These texture features are computed and analyzed for its role in identifying land cover types like water, urban and vegetation areas (Chamundeeswari et al., 2009).

Effect of three parameters, namely two roughness measures, 's' and 'l', and window size on 'mean' is studied.

2.3 Computing surface roughness from texture measures

Polynomial regression is employed on observations to express each of the texture measures in terms of window size and roughness measures, 's' and 'l'. Defining equations and its coefficients with R² values is given in Table 1.

It is observed from Table 1 that 'semivariogram', 'weighted rank fill ratio' and 'wavelet component 1' have higher impact on 's', 'l' and 'w' because R² value is higher than 0.9, whereas other texture measures like 'mean', 'variance' and 'lacunarity' have R² values less than 0.7. Hence, only the three texture measures, semivariogram, weighted rank fill ratio and wavelet component 1 are considered for further analysis.

Taking care of all these analysis, a relationship among RMS height 's', semivariogram, 'weighted rank fill ratio' and wavelet component 1 has been developed using regression analysis, that can be expressed as

$$s = b_0 + b_1 (\text{semivariogram}) + b_2 (\text{weighted rank fill ratio}) + b_3 (\text{wavelet component 1}) \quad (3)$$

By polynomial regression fit to the observed data (same set of data of approximately 1500 images used for computing texture measures in section 2.1), the coefficients b_0, b_1, b_2, b_3 are obtained as 0.0110, 2.3915, 5.4634 and -2.4103 respectively and R² is 0.9739. Similarly, polynomial regression fit is tried to obtain relationship for correlation length in terms of these measures, but R² is 0.4768 that confirms our earlier findings that the three texture measures namely 'semivariogram', 'weighted rank fill ratio' and 'wavelet component 1' are less sensitive on correlation length, 'l'. Figure 1 shows the scatterplot of retrieved RMS height with actual RMS height and it can be inferred that retrieved RMS height closely follows actual RMS height. This shows that surface roughness parameter, RMS height can be retrieved effectively using the mentioned three texture measures with the help of equation(3).

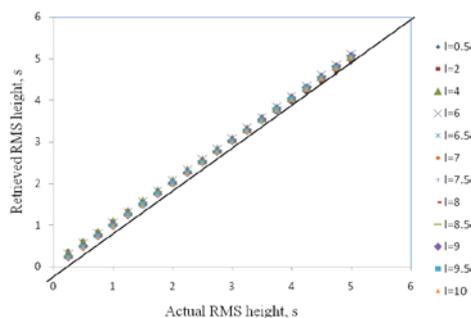


Figure 1. Scatter plot of retrieved RMS height vs actual RMS height, 's'.

III. Computation of roughness parameter, RMS height for real SAR image

3.1 Data used

ERS-2 SAR C-band image acquired on July 23, 2001 is used for developing and analyzing the proposed approach of unsupervised classification of single polarized SAR image. ERS-2 SAR C-band image of July 28, 2003 is used for validating the results for labeling. Solani river catchment around Roorkee town in the state of Uttarakhand, India is taken as the study area. The area is relatively flat with elevations ranging from 245.5 m to 289.9 m. Its latitude ranges from 29.90° N to 29.83° N, and its longitude ranges from 77.92° E to 77.85° E.

3.2 Proposed methodology to compute RMS height, 's' for SAR image

Clusters obtained by any of SAR segmentation methods can be taken as input for labeling. These clusters are to be labeled with major land cover types of water, urban and vegetation areas.

Steps involved in computing RMS height can be stated as listed

- In each of the cluster obtained by segmentation (Chamundeeswari et al., 2007; 2009), image is divided in to groups of pixels of size 5×5. For each of these pixel groups, three texture measures viz., 'semivariogram', 'weighted rank fill ratio' and 'wavelet component 1' are calculated.
- Using the equation (3), rms height, 's' is computed from these three texture measures.
- Similarly, RMS height is calculated for all groups of pixels of size 5×5 in each of the clusters.

Flow chart for computing RMS height 's' is given in figure 2.

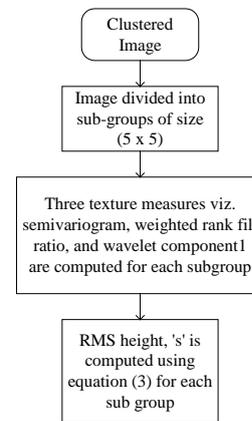


Figure 2. Flow chart for computing RMS height 's'

Critical analysis to label the various clusters

- Segmented image obtained by PCA based fusion approach (Chamundeeswari et al., 2007; 2009) is labeled with the help of topographic sheet and hence label of clusters are known.
- RMS height 's' is calculated for all groups of pixels within every cluster belonging to various land cover types like water, urban and agriculture areas. In Table 2, three texture measures, semivariogram, weighted rank fill ratio

and wavelet component 1 and RMS height 's' calculated from these measures for 10 samples of each land cover types, i.e., water, urban and vegetation areas are listed.

- From the calculated RMS height 's' for all groups of pixels with in each land cover type, range of 's' can be specified for that land cover. Similarly, range of 's' is computed for urban and vegetation areas and are shown in Figure 3.

Various values of texture measures with computed RMS height (from equation 3) is shown in Table 2. Typical ranges are presented in Table 3.

From the analysis, it is observed that RMS height 's' may be used to differentiate land cover on the basis of roughness. It is observed that with the help of RMS height 's' and backscattering coefficient, clusters can be labeled. If the RMS height 's' is <0.5 cm and backscattering coefficient is in the range of (-24 to -28)dB, then the cluster is labeled as water. Similarly, the conditions are applied for labeling urban and agriculture areas. The decision rules for labeling the clusters are tested with SAR data of July 28, 2003. SAR segmentation process is carried out to segment the images into various clusters. Now, the clusters are labeled with the proposed approach. For this purpose, pixels within each cluster are grouped into sets of 5×5 pixels. Then, each of the groups of pixels, texture measures namely 'semivariogram', 'weighted rank fill ratio' and 'wavelet component 1' are computed. Using equation 3, RMS height 's' is computed from these texture measures. On the basis of RMS height and the mean backscattering coefficient within each window, the groups of pixels are labeled as one of the major land cover types i.e., water, urban and agriculture areas (figure 3).

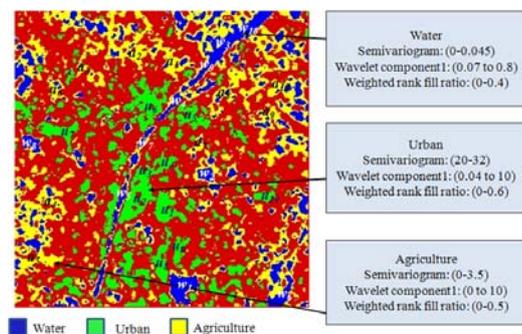


Figure 3. Clustered SAR image- texture measures for each cluster.

Conclusion

Texture measures like mean, variance, semivariogram, lacunarity, 'weighted rank fill ratio', wavelet components are analyzed critically for their variation when surface roughness measures, RMS height 's' and correlation length 'l' are varied. This is accomplished by developing synthetic images with varying 's' and 'l'. When computing texture measures, window size is varied to observe and incorporate the effect of window size, if required. From the analysis, it is found that three texture measures, semivariogram, 'weighted rank fill ratio' and 'wavelet component 1' are suitable for measuring 's'. It is found that 'l' and window size are less sensitive for these three texture measures, which helps to propose an empirical relation among 's' and these three texture measures. The empirical

relation developed with the synthetic images has been validated with real SAR images and results are quite encouraging. It means various major clusters like water, urban or agriculture areas can be easily labeled by considering roughness and backscattering coefficient. This infers that use of roughness with backscattering coefficient may resolve the problem of labeling of various clusters.

REFERENCES

- Anitha D., Mohan S., Patel A. B., 2006. Multi-frequency SAR signatures of forest class covering parts of Rajpipla, Gujarat. *Proceedings of SPIE, the International Society for Optical Engineering*, 6410, pp. 641005.1-641005.11.
- Chamundeeswari V. V., Singh D., Singh K., 2007. An adaptive method with integration of multi-wavelet based features for unsupervised classification of SAR images. *Journal of Geophysics and Engineering*, 4, pp. 384-393..
- Chamundeeswari V. V., Singh D., Singh K., 2009. An analysis of texture measures in PCA-based unsupervised classification of SAR images. *IEEE Geoscience and Remote Sensing Letters*, 6(2), pp. 214-218.
- Dobson M. C., Pierce L., Sarabandi K., Ulaby F. T., and Sharik T., 1992. Preliminary analysis of ERS-1 SAR for forest ecosystem studies. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2), pp. 203-211.
- Dobson M. C., Ulaby F. T., and Pierce L. E., 1995. Land-cover classification and estimation of terrain attributes using synthetic aperture radar. *Remote Sensing of Environment*, 51(1), pp. 199-214.
- Fung A. K., and Chen M. F., 1985. Numerical simulation of scattering from simple and composite random surfaces. *Journal of Optical Society of America, A-2*, 2(12), pp. 2274-2284.
- Pant, Singh D., and Srivastava T., 2010. The Potential Application of Fractal Approach for Surface Roughness Retrieval: A Study for Simulated Surfaces. *International Journal of Geomatics, Natural Hazards and risk Management (Published by Taylor & Francis)*, vol. 1, no.3, pp. 243-257.
- Prakash R., Singh D., 2008. Microwave sensitivity analysis of soil texture at C-band with bistatic scatterometer for remote sensing. In *International Conference on Recent Advances in Microwave Theory and Applications, MICROWAVE*, pp.211-213.
- Qiang Y., Fang C., and Wen H., 2007. A validation method for bare soil surface parameters inverted by Oh empirical model using integral equation method. *1st Asian and Pacific Conference on Synthetic Aperture Radar*, pp. 211-214.
- Teng H.T., Ewe H.T., and Tan S.L., 2010. Multi-fractal dimension and its geometrical terrain properties for classification of multi-band multi-polarised SAR image. *Progress in Electromagnetic Research*, 104, pp. 221-237.
- Xia Z-G., and Henderson F. M., 1997. Understanding the relationships between radar response patterns and the bio- and geophysical parameters of urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 35(1), pp. 93-101.

Table 1. Texture measures in terms of roughness parameters and window size.

a_0^{1-6} , a_1^{1-6} , a_2^{1-6} and a_3^{1-6} are the coefficients of defining equations for texture measures (sl. no. 1 to 6) and y_{1-6} represents six texture measures, 'mean', 'variance', 'semivariogram', 'lacunarity', 'weighted rank fill ratio' and 'wavelet component 1' respectively. 's', 'l' and 'w' represents RMS height, correlation length and window size respectively.

Sl. No	Textural measures	Defining equation	Parameters				R ²
			a_0^{1-6}	a_1^{1-6}	a_2^{1-6}	a_3^{1-6}	
1	Mean, y_1	$y_1 = a_0^1 + a_1^1 s + a_2^1 l + a_3^1 w$	-4.7766	5.2276	0.1980	0.3989	0.6942
2	Variance, y_2	$y_2 = a_0^2 + a_1^2 s + a_2^2 l + a_3^2 w$	-0.0152	0.1413	-0.0017	-0.0004	0.6775
3	Semivariogram, y_3	$y_3 = a_0^3 + a_1^3 s + a_2^3 l + a_3^3 w$	-0.0004	0.1694	-0.0042	0.0014	0.9145
4	Lacunarity, y_4	$y_4 = a_0^4 + a_1^4 s + a_2^4 l + a_3^4 w$	15.6121	-0.5422	7.2148	-1.3732	0.6521
5	Weighted rank fill ratio, y_5	$y_5 = a_0^5 + a_1^5 s + a_2^5 l + a_3^5 w$	-0.0616	0.882	0.032	-0.0001	0.9368
6	Wavelet component 1, y_6	$y_6 = a_0^6 + a_1^6 s + a_2^6 l + a_3^6 w$	-0.1258	1.7652	0.0659	0.0014	0.9365

Table 2. RMS height calculated in real SAR image for water, urban and agriculture areas.

(w_1-w_{10} represent 10 samples of water areas and corresponding texture measures with RMS height 's' are listed. Similarly u_1-u_{10} represent 10 samples of urban areas and a_1-a_{10} represent 10 samples of agriculture areas.)

Land cover type	Texture Measures			RMS height
	Semivariogram	Weighted rank fill ratio	Wavelet component 1	
Water [w_1-w_{10}]	[0.045,0.045, 0.034,0.023, 0.014,0.036, 0.033,0.018, 0.042,0.044]	[0.4,0.4, 0.34,0.37, 0.24,0.22, 0.16,0.07, 0.34,0.32]	[0.8,0.07, 0.4,0.53, 0.56,0.24, 0.45,0.7, 0.65,0.45]	[0.472,0.443, 0.398,0.347, 0.298,0.399, 0.390,0.32, 0.449,0.45]
Urban [u_1-u_{10}]	[30,28, 26,24, 22,21.8, 20,28, 24,27]	[0.6,0.57, 0.54,0.45, 0.38,0.33, 0.5,0.22, 0.14,0.05]	[10,8.7, 9.3,5.7, 6.3,6.3, 9.7,7.9, 3.8,4.2]	[156.37,145.94, 135.58,125.05, 114.68,113.64, 104.44,145.89, 124.96,140.55]
Agriculture [a_1-a_{10}]	[0.8,2.1, 2.7,3.2, 0.28,0.19, 0.43,0.38, 0.46,0.22]	[0.5,0.43, 0.36,0.48, 0.53,0.27, 0.63,0.17, 0.34,0.28]	[10,4, 5.1,6.3, 8.3,9.1, 8.2,7.6, 4.8,9.6]	[4.76, 11.27, 14.42,17.07, 1.99,1.55, 2.77,2.47, 2.78,1.72]

Table 3. Typical range of RMS height for major land covers.

Land cover type	Water	Urban	Agriculture
Typical values of RMS height 's' (cm)	<0.5	(100 and higher)	(0.7-20)