PROPOSAL OF A WEIGHTED INDEX FOR SEGMENTATION EVALUATION

Mariane Souza Reis, Eliana Pantaleão, Sidnei João de Siqueira Sant'Anna, Luciano Vieira Dutra

Image Processing Division - Brazilian National Institute for Space Research (INPE)
{reis, elianap, sidnei, dutra}@dpi.inpe.br

ABSTRACT

Although segmentation is an important process in image classification, selecting among different segmentors and their parameters is a difficult task. This work proposes a reference free index that returns the quality of segmentation, considering the classes that the user intends to classify. Considering a gaussian distribution, this index was tested to evaluate segmentations of an optical simulated image, a LANDSAT5/TM image in a Brazilian Amazon area and its derived fraction image. Index results presented higher values for segmentations more similar to the reference image, and also good agreement with overall accuracy values when classifying the images.

***Index Terms*—** Segmentation index, segmentation evaluation, optical data segmentation

1. INTRODUCTION

Segmentation plays an important role in remote sensing image classification. The pixelwise classification results often show some isolated pixels due to outliers, noise on the imaging process or single objects in the scene that are not relevant to the analysis. For the aim of producing land cover maps, object-based or region based classification is usually a better option to obtain more homogeneous regions.

There have been many attempts to evaluate segmentation results, as, for example, [1] and [2]. Most of them try to assess the quality of the segmentation regardless of its purpose. They are based only on the characteristics of the segments with respect to the original data set.

This research proposes an index that measures the suitability of a given segmentation to a specific classification task. The index has three components: I, v and $Dist$. I and v are unsupervised and assess the quality of the segmentation in general: I measures the homogeneity of the segments and v, the difference among the adjacent segments. The third component, $Dist$, relates the segmentation to a class set, and is used to weight the other parameters. With the use of this component, the index has higher (i.e. better) values if different classes are in different segments.

Funded by CNPq grants #150439/2014-8, 307666/2011-5.

2. PROPOSED INDEX

The proposed index is an improvement of the index developed by [3]. The aim of a segmentation is to maximize intra-segment homogeneity and intersegment heterogeneity. Both indices reflect these measures, representing the quality of a segmentation as a combination of intra-segment variance (v) and Moran's Index (I). The higher value represents the best segmentation in a given image.

In the objective function (OF) proposed in [3], the index is vulnerable to the number of segmentation attempts, because the normalization is made after the indices are calculated. The Objective Function proposed by [3] is given by:

$$OF = F(v) + F(I) \tag{1}$$

in which:

$$F(x) = \frac{X_{max} - X}{X_{max} - X_{min}} \tag{2}$$

$$v = \frac{\sum_{i=1}^{n} a_i.v_i}{\sum_{i=1}^{n} a_i} \tag{3}$$

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \overline{y})(y_j - \overline{y})}{(\sum_{i=1}^{n}(y_i - \overline{y})^2)(\sum_{i\neq j}\sum w_{ij})} \tag{4}$$

v_i is the variance and a_i is the area of region i; n is the total number of regions, y_i is the mean digital number of region R_i, y_j is the mean digital number of region R_j and $\overline{y}$ is the mean digital number of the image. w_{ij} is the measure of the spatial adjacency of regions R_i and R_j. If these regions are adjacent, w_{ij} is one. Otherwise, it is zero.

The OF measure is completely independent of the classification task. This independency means that a segmentation that does not distinguish two features that belong to different but similar classes may have a higher index value than one that separates those features. The use of a supervised component allows the adjustment of the segmentation to the desired class set. Besides, the OF measure was formulated for one band only.

To solve the problems found in the use of OF, we propose a weighted index for segmentation evaluation ($WISE$) that is normalized regardless of the number of segmentations tested.

While OF is calculated for each band, $WISE$ uses information from all image bands to calculate the index for a specific segmentation. The idea is to use the lowest value of the intra-segment component, because if an object is distinguishable in one band, but the segmentation does not separate this object, the index must return a low value. On the other hand, if the object is distinguishable in one band, it is an individual object and should be separated, regardless of its discrimination in other bands. Therefore, the highest value of I' among the bands is used, since it returns a low value in over-segmented images. $WISE$ is calculated by:

$$WISE = \frac{1}{Dist}.min(v') + max(I') \quad (5)$$

v' and I' are modified versions of v and I given by:

$$v' = 1 - \frac{v}{\sigma^2} \quad (6)$$

$$I' = 1 - |I| \quad (7)$$

σ^2 is the variance of the whole image and $|I|$ is the absolute value of I.

$Dist$ is the minimum distance (with distance as a measure of divergence between two elements) calculated between any pair of classes. In this work, it was used Jeffries-Matusita (JM) distance, that is given by:

$$JM_{ij} = \sqrt{2(1 - e^{-B_{ij}})} \quad (8)$$

in which B_{ij} is the Bhatthacharyya distance between the classes i and j. For gaussian distributions it is given by:

$$B_{ij} = \frac{1}{8}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2}\right)^{-1} (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + \frac{1}{2} ln \frac{\left|\frac{\Sigma_i + \Sigma_j}{2}\right|}{\sqrt{|\Sigma_i|\,|\Sigma_j|}} \quad (9)$$

μ_i is the mean of class i, μ_j is the mean of class j, Σ_i is the covariance matrix of class i and Σ_j is the covariance matrix of class j. If one does not desire to consider the classes in the analysis, it is possible to set $Dist$ to one, so that the components I' and v' have the same weight.

3. METHODOLOGY FOR INDEX EVALUATION

Performance of $WISE$ was assessed in two approachs. In the first one, the segmentations that obtained the higher values of $WISE$ were visually compared to a reference segmentation. In the second approach, the index value of different segmented images were compared to the overall accuracy (OA) index of the respective classification.

A simulated image having optical characteristics was used in both approaches. A remote sensed image and a fraction image were also used in the second approach.

The real image was a 458x333 pixels LANDSAT5/TM image (WRS 227/62, June 29th, 2010), bands 2, 3 and 5. The fraction image was formed by the vegetation, soil and shadow fractions of the TM image (calculated using bands 1-5 and 7 and the method proposed by [4]). These images cover part of the BR-163 (Cuiabá-Santarém Highway), in Belterra, Pará state, located in the Brazilian Amazon. The region presents humid tropical climate and dominant vegetation is Humid Tropical Rainforest [5]. As a result of the occupation process along the highway, there are mosaics of secondary vegetation in varying stages of development, pasture, croplands and bare soil areas within the forest matrix [6].

The method proposed by [7] was used to create the simulated image. An image model based on the mean vector and covariance matrix of some classes was extracted from a real image and applied to a phantom image (an idealized cartoon image containing the regions, in this work, a 496x496 pixels hand-drawn image containing 78 regions). Bands 3, 4 and 5 of the same LANDSAT5/TM image described were used to estimate the mean vector and covariance matrix of six different classes. These classes were: cultivated agricultural areas, primary forest, managed pasture, unmanaged pasture, bare soil and secondary vegetation. The image was simulated considering a Multivariate Gaussian distribution, differences up to 10% of mean and standard deviation among regions of the same class and no spatial correlation.

All the images were normalized to mean 127 and standard deviation 42 before segmentation. This values were empirically determined to improve segmentation when using region growing algorithm. The segmented images used in both approaches were created using different similarity values in Terraview's region growing based segmentor, to each image separately. The minimal area parameter was set to 20.

The classified images were created using the aforementioned segmentations and the Bhattacharyya distance classifier, as implemented by [7]. For the simulated image, the six original classes were considered. For TM and fraction images, ten cover classes were used to train the classifier: primary forest, degraded forest, managed pasture, unmanaged pasture, cultivated areas, bare soil, idle agricultural areas and secondary vegetation in initial, intermediate and advanced stages.

4. RESULTS

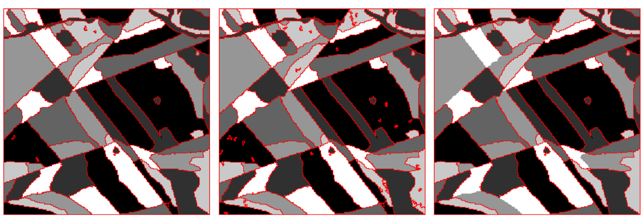
$WISE$ results for each segmentation are presented in Table 1. The notation for parameters used in each segmentation is "t_a", in which a is the similarity parameter.

In the simulated image, the best result of $WISE$ was obtained by t_30. Figure 1 presents this segmentation over the reference segmentation. The segmentations with immediate lower and higher similarity parameters are also presented, for comparison.

It is possible to observe that although t_30 has some small

Table 1. $WISE$ results for Simulated, TM and Fraction Images.

Segmentation	Simulated Image	TM Image	Fraction Image
t_05	1.051	2.406	2.716
t_10	1.057	2.418	2.805
t_15	1.131	2.424	2.838
t_20	1.247	2.374	2.825
t_25	1.442	2.411	2.825
t_30	1.756	2.391	2.854
t_35	1.557	2.382	2.817
t_40	1.435	2.143	2.700
t_45	1.419	2.115	2.605
t_50	1.419	2.035	2.533
t_55	1.419	2.023	2.479
t_60	1.341	2.031	2.348



(a) t_30 (b) t_25 (c) t_35

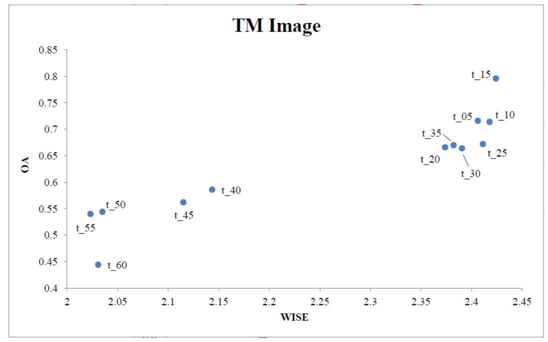
Fig. 1. Reference segmentation of simulated image with segmentations superposed.

polygons in non existing regions, t_35 has regions that overlaps two different classes, clustering two or three regions in one polygon. Obviously, t_25 shows more small polygons than t_30. As the evaluated classes were high separable ($Dist$ equal to 1.12), under-segmentations obtained higher values than over-segmentations.

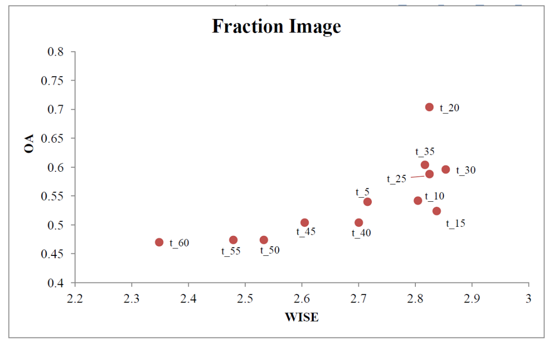
$WISE$ results and OA indexes for original TM, fraction and simulated images segmentations are presented in Figure 2. It can be seen that for TM and fraction images, higher values of $WISE$ tend to present higher OA values, although the classifiers seem to work better with over-segmentated images.

Figure 3 presents a subset of the original TM image and the segmentation with the higher $WISE$ value, as well as those with more similar parameters, for comparison. It is possible to observe that in t_20 some features of more a more homogeneous forest is not separated from the uneven one. This problem is solved in t_15, although the rest of the image is over-segmented.

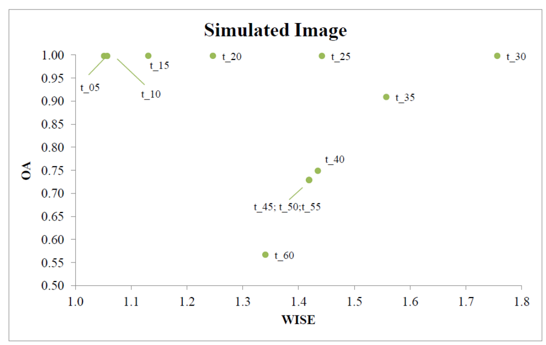
Figure 4 presents the segmentations for the whole fraction image, with the segmentations that obtained the higher $WISE$ value (t_30) and the higher OA value (t_20). While in t_30 some small features are not separated, in t_20 bigger features are over-segmented. As said before, the classifier is prepared to work with small segments, and this induced the best OA in t_20. $WISE$, as well as the OF proposed by [3],



(a)

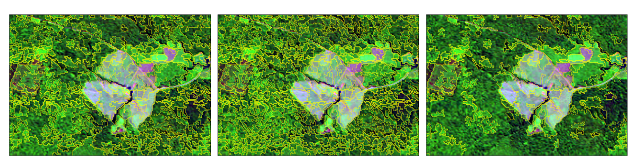


(b)



(c)

Fig. 2. $WISE$ and OA values comparison for classified images.

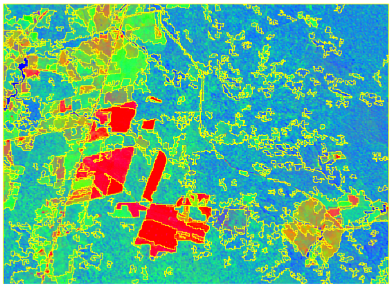


(a) t_15 (b) t_10 (c) t_20

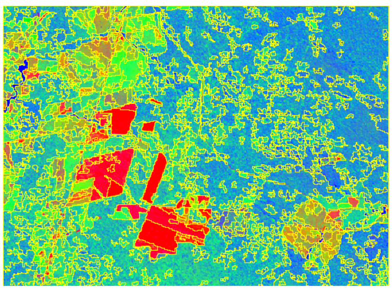
Fig. 3. Subset of the original TM image with segmentations superposed. Color composition: B5(R)B4(G)B2(B), 2% contrast.

attributes lower weigths to small regions in the index calculation, therefore atributing the higher value to t_30.

In the simulated image classification it is possible to ob-



(a) t_30



(b) t_20

Fig. 4. Fraction image with segmentations sobreposed. Color composition: Soil(R)Vegetation(G)Shadow(B), 2% contrast.

serve two different behaviors. The first one refers to the non-variant values for OA when varying the similarity parameter of segmentation from 5 to 30. The second one is the decreasing values of OA, and respective decreasing of $WISE$, with the increasing of the similarity parameter. This happens because of the relative high separability of the considered classes in simulated images. The classifier correctly separates the classes, even in small regions, so OA is high until the point when the segmentations themselves join different classes, introducing errors in the classification.

5. CONCLUSIONS AND CONSIDERATIONS

Results indicate a strong agreement of $WISE$ and the overall accuracies of each classifications. In this sense, it is possible to affirm that high $WISE$ values are associated with the best segmentations when using LANDSAT5/TM derived data and Terraview's region growing segmentor. Future works include testing $WISE$ with other images, including well known and simulated ones, and comparing the results with other segmentations indices. $WISE$ will be also evaluated for multi-level class sets.

That said, some aspects of $WISE$ must be considered. The first one is that when segmenting images with at least one pair of very similar classes, $WISE$ returns high values for over-segmented images. However, this result can only be considered the best if the user intends to separate these specific classes, in the sense that they probably share at leat one border, otherwise there will be no reason for over-segmentation. Therefore, the user must choose the minimal distance JM between pairs of classes that do share borders. If this information is unknown, a reliable alternative is to set $Dist$ to 1.0, so that I' and v' has the same weight. Another aspect to be considered is that if the user intends to separate two classes that have JM distance near to zero, the index will recommend the segmentation most similar to the pixels themselves. Also, there is the mathematical impossibility to calculate the index when $Dist$ equals to zero.

At this stage, the index is being tested with another segmentors, in order to further verify its efficiency. Also, $WISE$ using another measures of stochastic distances to weigth the I' and v are being tested for Synthetic Aberture Radar images.

6. REFERENCES

[1] Hui Zhang, Jason E. Fritts, and Sally A. Goldman, "Image segmentation evaluation," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 260–280, 2008.

[2] P.R Marpu, M. Neubert, H. Herold, and I. Niemeyer, "Enhanced evaluation of image segmentation results," *Journal of Spatial Science*, vol. 55, no. 1, pp. 55–68, 2010.

[3] G. M. Espindola, G. Câmara, I.A. Reis, L.S. Bins, and A.M. Monteiro, "Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation," *International Journal of Remote Sensing*, vol. 27, no. 14, pp. 3035–3040, 2006.

[4] Y.E. Shimabukuro and A. Smith, "The least-squares mixing models to generate fraction images derived from remote sensing multispectral data," *IEEE Transactions on Geosciense and Remote Sensing*, vol. 29, no. 1, pp. 16–20, 1991.

[5] Brazil, "Floresta nacional do Tapajós: Plano de manejo," Report, volume i (in portuguese), Brazilian Institute of Environment and Renewable Natural Resources (IBAMA), 2004.

[6] M. Escada, S. Amaral, C.D. Rennó, and T.F. Pinheiro, "Levantamento do uso e cobertura da terra e da rede de infraestrutura no distrito florestal da BR-163," Report (in portuguese), Instituto Nacional de Pesquisas Espaciais (INPE), 2009.

[7] R.G. Negri, L.V. Dutra, and S.J.S. Sant'Anna, "Stochastic approachs of minimum distance method for region based classification," *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Aplications Lecture Notes in Computer Science*, vol. 7441, pp. 797–804, 2012.