

A Geographical Approach to Self-Organizing Maps Algorithm Applied to Image Segmentation

Thales Sehn Korting, Leila Maria Garcia Fonseca, and Gilberto Câmara

Image Processing Division
National Institute for Space Research – INPE
São José dos Campos – SP, Brazil
{tkorting,leila,gilberto}@dpi.inpe.br

Abstract. Image segmentation is one of the most challenging steps in image processing. Its results are used by many other tasks regarding information extraction from images. In remote sensing, segmentation generates regions according to found targets in a satellite image, like roofs, streets, trees, vegetation, agricultural crops, or deforested areas. Such regions differentiate land uses by classification algorithms. In this paper we investigate a way to perform segmentation using a strategy to classify and merge spectrally and spatially similar pixels. For this purpose we use a geographical extension of the Self-Organizing Maps (SOM) algorithm, which exploits the spatial correlation among near pixels. The neurons in the SOM will cluster the objects found in the image, and such objects will define the image segments.

1 Introduction

Image segmentation is currently one of the most challenging tasks in digital image processing. The results of segmentation are used by many other tasks regarding information extraction from images. One simple definition states that a good segmentation should separate the image into simple regions with homogeneous behavior [7].

Algorithms for segmentation covers splitting one image into its components, concerning a specific context. The context includes scale, since regions from segmentation can represent different levels of perception. In an urban remote sensing image, found objects describe buildings in a finer scale, or city blocks in a coarser scale. Depending on the application, both scales provide enough information for object detection. Context also includes neighborhood, in accord to Tobler's law, which states that near things are more related than distant things [18]. In the image processing area, this law may be interpreted as *near pixels are more similar than distant pixels*. Extending this interpretation to image segmentation, we can state that near and similar pixels should describe regions in the segmentation results.

There are several applications in remote sensing which are depending on the results of image segmentation. The found objects must be in accord to the targets of a satellite image, such as roofs [5], streets [8] and trees in urban areas. In

agricultural imagery, the algorithm should extract different crops [16], or deforested areas in forests [17]. Such different land covers differentiate land uses by classification algorithms.

Several initiatives of segmentation algorithms applied to satellite imagery have already been proposed in the literature. However, none of them exploits deeply the spatial correlation among points as suggested by Tobler's Law. In this scenario we present a novel technique for segmentation which aims to fulfill this gap. The paper is organized as follows. Section 2 presents the literature review, Section 3 presents the methodology, Section 4 shows the results and the Section 5 concludes the paper.

2 Literature Review

Traditional segmentation algorithms are often based on the pixel similarity, by comparing one pixel to its neighbors. Several strategies are based on the technique named "region growing", which merges neighboring pixels with a certain degree of similarity. Some examples of algorithms are defined in [4] and [1]. The similarity are often based only on spectral properties, and seldom refers to the location of the pixels.

Many other strategies for image segmentation have already been proposed in the literature. For further reading on traditional segmentation techniques, please refer to [19]. The present work integrates one classification approach to perform segmentation. The employed classification algorithm is a modification of the well known Self-Organizing Maps – SOM [10]. This work investigates an approach to merge spectrally and spatially similar pixels by using a geographical extension of SOM algorithm, named GeoSOM [2].

2.1 SOM Applied to Remote Sensing

SOM is a powerful tool for exploring huge amounts of high-dimensional data. It defines an elastic, topology-preserving grid of points that fits to the input space [12]. Initially proposed as a visualization tool for n-dimensional data, SOM have been applied in a variety of applications, such as clustering, dimensionality reduction, classification, sampling, vector quantization, and data mining [10].

Basically, according to [21], the SOM training algorithm is similar to a K-means clustering, with an added spatial smoothness constraint. Neurons are initialized randomly, and at each iteration the neuron with most similar response to the input data is called winner. The winner and its neighbors are updated to become more similar to the corresponding input value. The size of the considered neighborhood is decreased along the epochs, so that on latter phases only the winner neurons are updated, with very small refinements.

In remote sensing several works employ SOM for classification issues. The work of [21] tested the algorithm to classify different vegetation types using radar images of agricultural areas. The idea of semi-supervised SOM was applied, where the closest neuron to a certain pattern is associated with one class

of the training objects. [11] performed a comparison between the Expectation-Maximization and SOM algorithms applied to intra-urban classification using high spatial resolution images to derive the results. According to the authors, both algorithms achieved good results, and presented their advantages. However, SOM classified the data in less time, which for the current increasingly amount of data, is an important advantage.

[9] have concluded that SOM is a good choice of an algorithm for monitoring land cover change. They investigated the optimal classification algorithm applied into multi-temporal data based on monthly phenological characteristics. [13] employed subpixel analysis using Landsat ETM+ data to estimate impervious surface coverage, lawn, and woody tree cover in typical urban landscapes. The authors combined SOM, Learning Vector Quantization (LVQ), and Gaussian Mixture Model (GMM) methods to estimate the posterior probability of the land cover components. [20] have applied SOM in the classification of multi and hyperspectral satellite imagery. Since SOM classification is considered to be a topology preserving projection of high-dimensional data onto a low-dimensional lattice, results emphasized the necessity of a faithful topological mapping for correct interpretation.

2.2 Segmentation Using SOM

Classifying image pixels to perform segmentation is not a new topic. [15] have presented one approach using hue and saturation image components to segment personal photographs. SOM was employed to estimate the main found chromaticities. Then, each pixel was classified according to the identified classes. Neighboring pixels relying in the same class were merged in the resultant segmentation. Also applied to personal photographs, [3] proposed one extension to the traditional SOM algorithm, which tried to map classes of pixel intensities in a hierarchical structure of neurons.

More recently, [12] have applied SOM in the segmentation of images for content-based image retrieval (CBIR). CBIR has been targeted at interactive use where the task is to return interesting or relevant images from an unannotated database. In their method, the color and texture values are used to train the map and to classify the segments. SOM have also been applied into magnetic resonance brain image segmentation [14], including spatial constraints using a Markov Random Field (MRF) model to improve the results.

3 Methodology

The main idea is to employ classification to perform segmentation, *i.e.* neighboring pixels with the same class are merged into segments. The novelty is to take advantage of the spatial correlation among the pixels, since we adopted a geographical extension of SOM algorithm, which considers the spatial proximity of the pixel elements when merging them into a single object. The SOM extension was proposed by [2] and is named GeoSOM. According to the authors, GeoSOM has the potential to organize the SOM output space according to the geographic proximities of the input patterns.

As GeoSOM finds patterns in the image pixels, it must fit homogeneous regions in these patterns, and considers these regions as the resultant segments. A simple way of inserting spatial constraints in classification is to include spatial-relevant variables which are computed as any other variable [2]. In this case remains the choice of what spatial variables to use. The possibilities are endless, and depend on the objectives pursued. Also one has to find the weights to be attributed to the geographic variables, thus giving more (or less) importance to geographic information.

The weights of the neurons have their own geographical properties. Since neurons have spatial positions, they are expected to converge to patterns found nearby in the image. In the resultant segmentation clusters reflect the spectral distribution of the data, and are positioned according the spatial occurrence of the patterns. Figure 1 shows this approach. In the following we describe the GeoSOM algorithm. Let:

- X be the set of N training patterns x_1, x_2, \dots, x_N , each of these having a set of components geo, i and another set $notgeo, i$;
 - W be a $p \times q$ grid of units w_{ij} where i and j are their coordinates on that grid, each of these units having a set of components $w_{geo, ij}$ and another set $w_{notgeo, ij}$;
 - α be the learning rate, assuming values in the interval $(0, 1)$;
 - $h(w_{ij}, w_{mn}, r)$ be a neighborhood function, used to update the neurons w_{ij} and w_{mn} with radius r , generally expressed by the gaussian equation
$$e^{-\frac{1}{2} \frac{(i-m)^2 + (j-n)^2}{r^2}}$$
 - k be the radius of the geographical Best Matching Unit (BMU) to be searched;
 - f be true if the units are at fixed geographical locations.
- 1 repeat
 - 2 for $m = 1 \rightarrow N$
 - 3 $\forall w_{ij} \in W$, calculate $d_{ij} = ||geo_m - w_{geo, ij}||$
 - 4 $w_{winnergeo}$ is the unit that minimizes d_{ij}
 - 5 select set $W_{winners}$ of w_{ij} that
$$||w_{winnergeo} - w_{ij}|| \leq k$$
 - 6 $\forall w_{ij} \in W_{winner}$, calculate $d_{ij} = ||x_m - w_{ij}||$
 - 7 w_{winner} is the unit that minimizes d_{ij}
 - 8 if $f = \text{true}$
 - 9 update units $w_{ij} \in W$: $w_{notgeo, ij} = w_{notgeo, ij} + \alpha h(w_{notgeo, winner}, w_{notgeo, ij}, r) ||x_m - w_{ij}||$
 - 10 else
 - 11 update units $w_{ij} \in W$: $w_{ij} = w_{ij} + \alpha h(w_{winner}, w_{ij}, r) ||x_m - w_{ij}||$
 - end if
 - 12 decrease α and r
 - 13 until $\alpha = 0$

During the training stage the algorithm searches for the nearest neuron to the input value, the so called $w_{winnergeo}$. In this case, “nearest” considers the

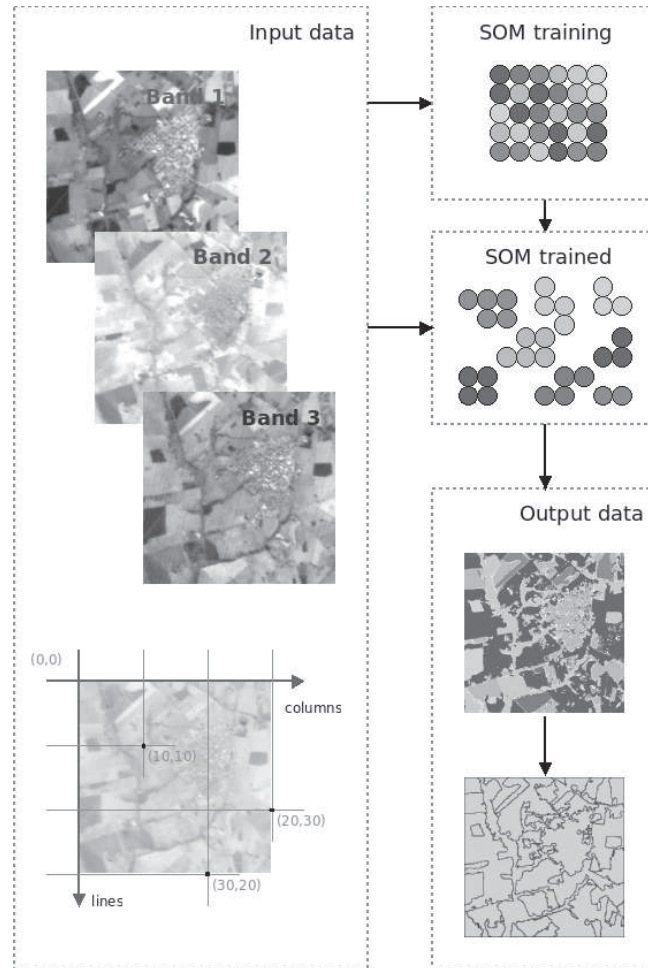


Fig. 1. Inserting spatial variables as input data – the GeoSOM approach

locations of the input and the neurons. After, it searches for the most similar neuron, in terms of spectral properties, within the neighborhood of $w_{winnergeo}$, finding the neuron w_{winner} . At this point the parameter f plays an important role. If f is true, the neurons are updated only in their spectral properties. If f is false, besides updating the spectral properties of the neurons, their location is changed as well.

The learning rate α and the radius r decrease along the epochs, allowing the algorithm to converge smoothly to the discovered patterns. The algorithm stops when one of these two parameters reaches 0, or when a pre-defined number of epochs is performed. Found patterns are then converted into segmentation. When the algorithm stops, neurons have their spectral properties similar to the

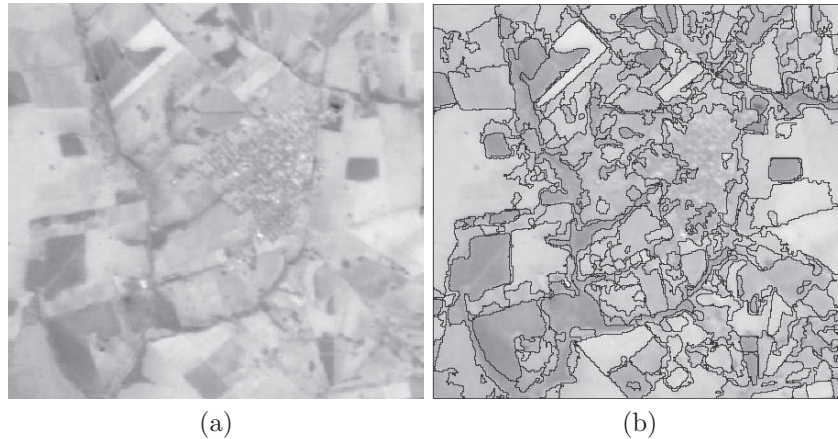


Fig. 2. Segmentation result: a) Input image, and b) segmentation with GeoSOM

main elements found in the image. Summarizing, the neurons in the SOM will cluster the objects found in the image, and such objects will define the image segments.

4 Results

The following results are composed by input images and their segmentation using GeoSOM. To produce such results we tuned the algorithm with a set of parameters that reached the best results, considering visual inspection. Nevertheless, we selected small image croppings and segmented them with different parameters. According to [6], “a strong and experienced evaluator of segmentation techniques is the human eye/brain combination”, and according [1], “no segmentation result – even if quantitatively proofed – will convince, if it does not satisfy the human eye”.

The first image is a crop of a region in the state of Minas Gerais, Brazil, using satellite CBERS-2B, instrument CCD¹. The spatial resolution of the image is 20m. Figure 2 shows the segmentation.

In the first result we used a map with 9 neurons (a 3 × 3 grid), with a learning rate $\alpha = 0.2$ for 50 epochs. 410 regions were generated. Segmented regions describe properly the main objects found in the image. In the central region, which is spectrally heterogeneous, the algorithm also obtained proper segments. As a drawback, some mistakes were noticed in the borders, as they were segmented as new objects. However, the training step in this example used merely 10% of the pixel information present in the image. This fact provides to our approach a significant reduction in the volume of data to be processed, reducing the processing time as well.

¹ Free remote sensing imagery at <http://www.dgi.inpe.br/CDSR/>

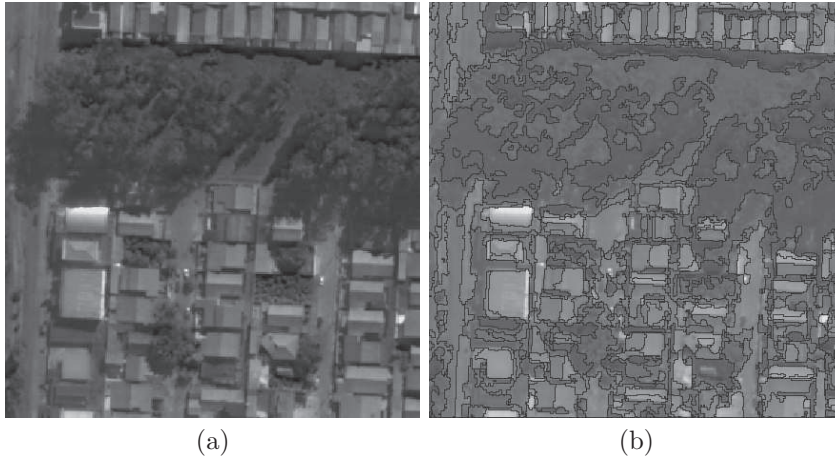


Fig. 3. Segmentation result: a) Input image, and b) segmentation with GeoSOM

We performed a second test using a high spatial resolution input image, segmenting a small crop of a Quickbird scene from an urban region in the state of São Paulo, Brazil. The spatial resolution of the image is $0.5m$. Target objects in this image are, differently from the first example, intra-urban objects, such as roofs, trees, shadows, streets, and so on. Figure 3 shows the segmentation.

In the second result we used a map with 25 neurons (a 5×5 grid), with learning rate $\alpha = 0.2$ for 200 epochs. 911 regions were generated. Since we used a higher amount of neurons, the algorithm generated more regions. Even the roofs of the houses were split, in some cases, in more than one region. By inspecting the values of the neurons responsible for the roofs regions, they presented similar spectral values, however different spatial locations.

The result shown in Figure 3 are considered over-segmentation, because in some cases the same object reflects in more than one segmented region. However it remains an useful result, since no region contains more than one object inside. In this case the main objects of the image can still be distinguished correctly, by applying proper inferences into the data.

5 Conclusion

This article presented the idea to apply an extension of the SOM algorithm, called GeoSOM, to remote sensing image segmentation. We described the main steps of the algorithm, and how to perform the segmentation. To evaluate the technique, we tested two images with different spatial resolutions and contexts. By tuning the parameters properly, we were able to generate visually appealing results, which makes this algorithm for segmentation encouraging approach.

One important aspect of this approach is the reduced amount of data required to perform the segmentation. With 10% of the total pixels in the images, the neurons were trained and the results were able to split properly the images

into their component regions. One drawback was detected in the segmentation of some borders, since they do not present spectral properties similar to the bordering objects. However, the smoothness of the resultant objects can be improved by adding a parameter to define a minimum area, or some other threshold applied to spatial properties of the generated segments.

Future research include to extend this approach by inspecting the dynamic of the neurons location along training. We believe that it is possible to describe the position of the most relevant objects in the image. If the image is classified after the segmentation, it is possible to infer the spatial distribution of the patterns. An interesting visualization tool can be derived from this approach, by retrieving the position and the amount of the neurons at each place of the image.

References

1. Baatz, M., Schape, A.: Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. In: Wichmann-Verlag (ed.) XII Angewandte Geographische Informationsverarbeitung, Herbert Wichmann Verlag, Heidelberg (2000)
2. Bação, F., Lobo, V., Painho, M.: Applications of different self-organising map variants to geographical information science problems. *Self-organising Maps: Applications in Geographic Information Science* (2008)
3. Bhandarkar, S., Koh, J., Suk, M.: Multiscale image segmentation using a hierarchical self-organizing map. *Neurocomputing* 14(3), 241–272 (1997)
4. Bins, L., Fonseca, L., Erthal, G., Li, F.: Satellite imagery segmentation: a region growing approach. In: *Brazilian Remote Sensing Symposium*, vol. 8 (1996)
5. Chesnel, A.L., Binet, R., Wald, L.: Object oriented assessment of damage due to natural disaster using very high resolution images. In: *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2007*, pp. 3736–3739 (2007)
6. Gamanya, R., Demaeyer, P., Dedapper, M.: An automated satellite image classification design using object-oriented segmentation algorithms: A move towards standardization. *Expert Systems with Applications* 32(2), 616–624 (2007)
7. Haralick, R., Shapiro, L.: Image segmentation techniques. *Applications of Artificial Intelligence II* 548, 2–9 (1985)
8. He, Y., Wang, H., Zhang, B.: Color based road detection in urban traffic scenes. In: *Proceedings of IEEE Intelligent Transportation Systems*, vol. 1 (2003)
9. Kim, D., Jeong, S., Park, C.: Comparison of Three Land Cover Classification Algorithms-ISODATA, SMA, and SOM-for the Monitoring of North Korea with MODIS Multi-temporal Data. *Korean Journal of Remote Sensing* 23(3), 181–188 (2007)
10. Kohonen, T.: *Self-Organizing Maps*, 3rd edn. Springer, Berlin (2001)
11. Korting, T.S., Fonseca, L.M.G., Bação, F.: Expectation-Maximization x Self-Organizing Maps for Image classification. In: *IEEE International Conference on Signal Image Technology and Internet Based Systems, SITIS 2008*, pp. 359–365 (2008)
12. Laaksonen, J., Viitaniemi, V., Koskela, M.: Application of Self-Organizing Maps and automatic image segmentation to 101 object categories database. In: *Proc. Fourth International Workshop on Content-Based Multimedia Indexing (CBMI 2005)*, Riga, Latvia, Citeseer (2005)

13. Lee, S., Lathrop, R.: Subpixel analysis of Landsat ETM+ Using Self-Organizing Map (SOM) Neural Networks for Urban Land Cover Characterization. *IEEE Transactions on Geoscience and Remote Sensing* 44(6), 1642–1654 (2006)
14. Li, Y., Chi, Z.: MR Brain image segmentation based on self-organizing map network. *International Journal of Information Technology* 11(8), 45–53 (2005)
15. Moreira, J., Costa, L.F.: Neural-based color image segmentation and classification using self-organizing maps. *Proceedings of the IX SIBGRAPI* 12(6), 47–54 (1996)
16. Perez, A., Benlloch, J., Lopez, F., Christensen, S.: Colour and shape analysis techniques for weed detection in cereal fields. *Computers and Electronics in Agriculture* 25, 197–212 (2000)
17. Silva, M., Câmara, G., Souza, R., Valeriano, D., Escada, M.: Mining patterns of change in remote sensing image databases. In: *The Fifth IEEE International Conference on Data Mining*, New Orleans, Louisiana, USA. Citeseer (2005)
18. Tobler, W.: A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46, 234–240 (1970)
19. Vergés-Llahí, J.: Color Constancy and Image Segmentation Techniques for Applications to Mobile Robotics. Ph.D. thesis, UPC (2005)
20. Villmann, T., Merényi, E.: Extensions and modifications of the Kohonen-SOM and applications in remote sensing image analysis. *Studies in Fuzziness and Soft Computing* 78, 121–144 (2002)
21. Wehrens, R.: Self-organising Maps for Image Segmentation. In: *Advances in Data Analysis, Data Handling and Business Intelligence: Proceedings of the 32nd Annual Conference of the Gesellschaft Für Klassifikation EV, Joint Conference with the British Classification Society (BCS) and the Dutch/Flemish Classification*, p. 373. Springer, Heidelberg (2009)