High Resolution Satellite Classification with Graph Cut Algorithms

Adrian A. López and José A. Malpica

Mathematics Department, School of Geodesy and Cartography, Alcalá University, 28871 Madrid, Spain
adrianl_98@yahoo.com, josea.malpica@uah.es

Abstract. In this paper, an unsupervised classification technique is proposed for high resolution satellite imagery. The approach uses graph cuts to improve the k-means algorithm, as graph cuts introduce spatial domain information of the image that is lacking in the k-means. High resolution satellite imagery, IKONOS, and SPOT-5 have been evaluated by the proposed method, showing that graph cuts improve k-means results, which in turn show coherent and continually spatial cluster regions that could be useful for cartographic classification.

Keywords: Graph cuts, k-means, high resolution imagery, unsupervised classification.

1 Introduction

Earth surface inspection by remote sensing provides valuable information not easily acquired by field survey techniques. Airborne and spaceborne images can assist in locating objects in the terrain and monitoring changes due to natural phenomena such as seasonal cycles, geomorphologic forces, or human activities. Multispectral remote sensing imagery is frequently used for land-cover and land-use determination.

Classification has been developed into a main technique to analyze a remotely sensed imagery, which represents a way to abstract information from it. There are two types of classification: unsupervised and supervised. Unsupervised classification does not require initial image area knowledge, while supervised classification uses training samples from image interest areas.

Unsupervised classification is applied to cluster images in order to extract representative classes and information. Usually, this information is employed to select accurate training samples to perform a supervised classification [1].

Today's algorithms, applied to low resolution satellite imagery, are not applicable high resolution satellites due to more complex and noisier spectral signatures for the latter. Therefore, new algorithms are needed to treat high resolution imagery.

The common unsupervised classification algorithm, when applied to imagery, generally does not consider the imagery's spatial characteristics data. Unsupervised classification in remote sensing is commonly applied, evaluating just the spectral information on a pixel by pixel basis. However, in images, pixels close together are more related than those which are far apart; therefore, information taken from neighboring pixels to those in the study provides a useful supplementary information

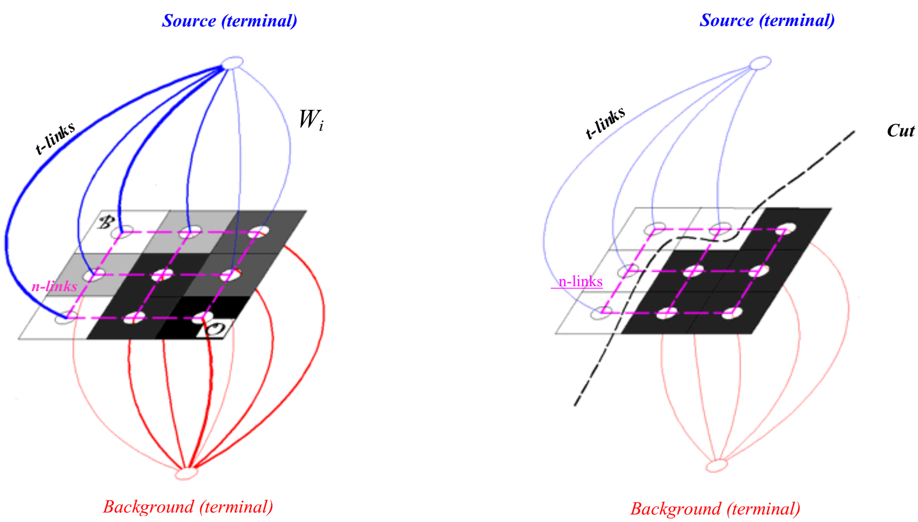
G. Bebis et al. (Eds.): ISVC 2008, Part II, LNCS 5359, pp. 105–112, 2008.

source. The idea of including the spatial domain of an image in the classification's process has been developed employing several techniques, including variogram analysis [2], texture [3], smart kernel operators [4], and wavelets [5]. Our research attempts to tackle the problem with graph cuts, which allow the introduction of spatial domain information into the classification. In the next section, we provide the basic definition and terminology of graph cuts and how they are used for image segmentation. Section 3 will present our research methods, and in Section 4 we present and discuss the results. Finally, our conclusion and the future direction of research are given.

2 Image Energy Segmentation with Graph Cuts

In this section, we provide basic terminology pertaining to graph cuts. In an undirected graph, $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, where $\mathcal{V}$ denotes the vertices and $\mathcal{E}$ represents the edges that connect these vertices. An example can be seen in Fig. 1; based on formula 1. Each edge $i \in E$ is assigned a non-negative weight (w_i). There are two special vertices called the terminals (source and background). A cut C of $\mathcal{G}$ is a partition of the vertices $\mathcal{V}$ into two sets, such that the terminals belong to each set. In combinatorial optimization, the cost of a cut is usually defined as the sum of the weights corresponding to the edges it serves [6] as is formulated in equation 1.

$$|C| = \sum_{i \in E} w_i \tag{1}$$



(a) (b)

Fig. 1. Graph cuts considering a two-dimensional image (3x3 pixels)

Graph vertices represent pixels (the white circles in the center of each square), and the edges represent relationships between neighborhood pixels, called n-links (represented by the dotted magenta lines in Fig. 1). There are also edges that flow from the terminals to the vertices (pixels) called t-links (represented by the red and blue lines in Fig. 1). The edges' thickness represents the weight w_i as can also be seen in Fig. 1.

The nine cells in the square of Fig. 1(a) represent a 3x3 image, where the grey level stands for a thresholding of the image on five levels. In Fig. 1(b), a graph cut segmentation is shown, which divides the image into two regions one for the object of interest (with the white label) and another for the background (with the black label).

To illustrate this type of phenomena, an energy is defined; then, energy minimization problems can be reduced to instances of the maximum flow problem in a graph [7], [8]. The "max-flow min-cut" theorem states that the maximum flow is equal to a minimal cut in capacity [9][10]. For example, in a network the maximum flow is dictated by its bottleneck (this is obtained with a cut on the graph as can be seen in Fig.1). Under most formulations of such problems in computer vision, the minimum energy solution corresponds to the maximum *a posteriori* estimate of a solution.

A globally minimized cut of a graph is a combinatorial optimization problem that can be computed in polynomial time. Greig et al. [11] were the first to apply graph cuts to image segmentation. This work remained unnoticed until the late 1990's, when various research employing graph cuts started to appear. Since then, these techniques have been employed efficiently in computer vision fields and image analysis to solve a wide variety of low-level computer vision problems, such as the stereo correspondence problem [12], multi-camera scene reconstruction [13], medical imaging [14], and many other computer vision problems that can be formulated in terms of energy minimization.

In cartographic segmentations, the objective is to group image pixels into logical areas that may represent cartographic interest objects on the terrain; this is the goal of our research, and will be further described in the next section.

Spatial domain information, or the neighborhood pixels relationship, is provided with constraints in the graph cut approach. These constraints are represented as labeling relationships within an energy function given by:

$$E(L)=\underbrace{\sum_{p\in P} D_p\left(L_p\right)}_{*}+\underbrace{\sum_{(p,q)\in N} V_{p.q}\left(L_p,L_q\right)}_{**}, \tag{2}$$

where $L=\left\{L_p \,\middle|\, p\in P\right\}$ represents image labeling P, D_p is the data penalty function, $V_{p.q}$ is an interaction potential, and *N* is a set of all pairs of neighboring pixels. The first term of Equation 2 is the regional term, and the second represents the boundary term.

Constrains are defined by the cost function as in Boykov and Kolomogorov's definition [6], where the boundary term is:

$$\sum_{\{p,q\}\in N} V_{p,q}(L_p,L_q)=\sum_{\{p,q\}\in N} B_{p,q}\cdot\delta_{L_p\neq L_q} ,$$

where

$$\delta_{L_p \neq L_q} = \begin{Bmatrix} 1 & \text{if} & L_p \neq L_q \\ 0 & \text{if} & L_p = L_q \end{Bmatrix}.$$

The boundary term (**) defines the costs of n-links, while the regional term (*) defines the costs of t-links. Inexpensive edges are attractive choices for the minimum cost cut. Hard constraints are implemented via infinity cost t-links. A globally optimal segmentation satisfying hard constraints can be computed efficiently in low-order polynomial time using max-flow/min-cut algorithms on graphs; the details can be seen in Ford and Fulkerson [10] and Goldberg and Tarjan [15].

We have shown the basic theory of graph cuts, considering two clusters and corresponding to the two terminals (source and background). Several classification

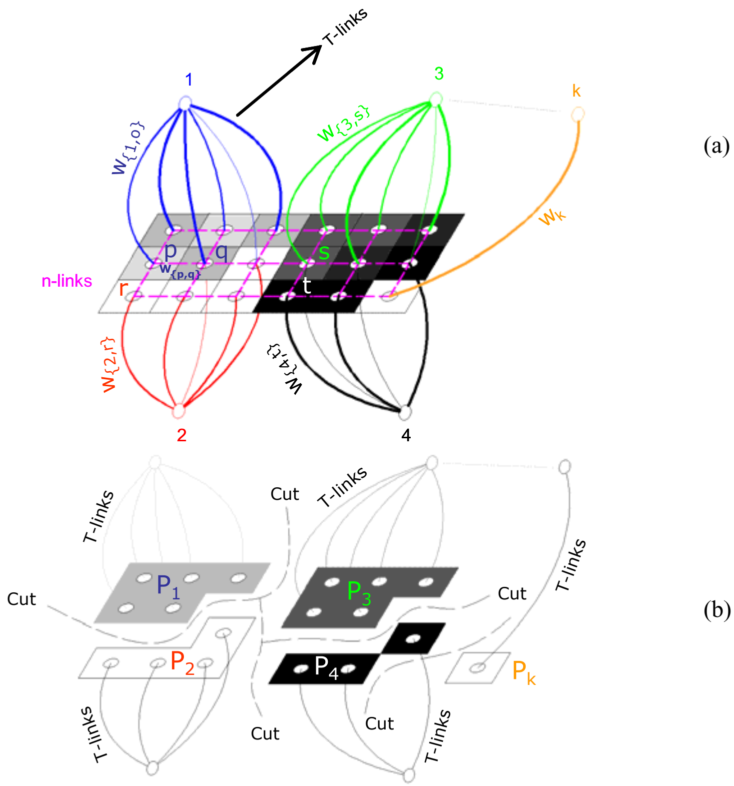


Fig. 2. Generalization to a multi-object. (a) A Graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, with multiple terminals 1,….., k (b) A multi-way cut on $\mathcal{G}$. The vertices (pixels) are shown as white circles. Each vertex has an n-link to its four neighbors, and is also connected to all terminals by t-links. The set of vertices $\mathcal{V}$ includes all pixels and terminals. The set of edges consists of all n-links and t-links. The induced graph $\mathcal{G}(\mathcal{C}) = \langle \mathcal{V}, \mathcal{E}-\mathcal{C} \rangle$ corresponding to some multi-way cut $\mathcal{C}$. A multi-way cut corresponds to a unique partition (labeling) of image pixels [16].

problems attempt to assign more than two labels to segmentation regions. We follow this generalization to a multi-object extraction problem proposed by Boykov and Funka-Lea [6].

3 Proposed Technique

We have used the flexible formulation for image segmentation provided by graph cuts to improve the k-means clustering algorithm on satellite images. Our technique is based on Bagon Shai's segmentation method [17]; we have adapted it to work on multiband satellite images.

The proposed technique has been applied to Ikonos and Spot-5 multispectral satellite imagery. These satellites obtain information in four bands (R-Red, G-Green, B-Blue, NR-Near Infrared for Ikonos; and R-Red, G-Green, NR-Near Infrared, MIR-Mid Infra Red for Spot-5), with spatial resolutions of 4 m (Ikonos), and 10-20 m (Spot-5), respectively (See Table 1).

Table 1. Ikonos and SPOT-5 spectral and spatial information

Sensor	Example	Pixel Size	Spectral Bands
IKONOS	1 Blue	4m	0.445-0.516 μm
	2 Green	4m	0.506-0.595 μm
	3 Red	4m	0.632-0.698 μm
	4 Near Infra Red	4m	0.757-0.853 μm
SPOT 5	1 Green	10m	0.50 - 0.59 μm
	2 Red	10m	0.61 - 0.68 μm
	3 Near Infra Red	10m	0.78 - 0.89 μm
	4 Mid Infra Red	20m	1.58 - 1.75 μm

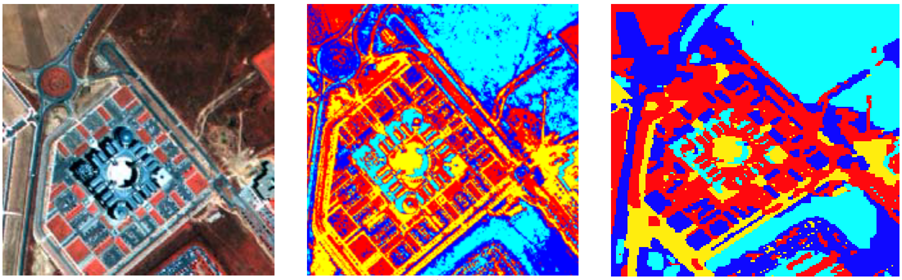
First, a k-means classification [18] is performed in the imagery. Let us assume that k classes are set, and the imagery has ***m*** bands. Then an *m-dimensional* Euclidean space is defined such as each pixel is a point (or vector) belonging to this space. The k-means groups pixels in clusters. The shift covariance matrix of the means (o centroid) of these clusters is the regional term. The reason the shift covariance matrix is used is because it gives a high weight to the pixels close to its centroid, and a low weight to the pixels situated far from its centroid; the further away the pixel is from its centroid, the lower its weight. For instance, if the labeled pixel belonging to a cluster is not strong enough, a high covariance will be assigned. This will be interpreted by the graph cut algorithm as a penalty.

A Sobel filter is applied to evaluate the boundary strength of the image region, in order to determine if a pixel is inside or outside the interest region. This is considered the boundary term.

There is an interesting relationship between graph cut and Markov Random Fields (MRF) examined by Boykov et al. [16]. Some low-level vision problems can be expressed in a Bayesian framework using MRF. Boykov et al. probed the equivalence of a specific class of MRF which generalizes the Potts model with the graph cut given a specific energy function. Recently, other authors such as del-Toro-Almenares et al. [19] have applied the graph cuts approach to MRF cartographic feature extraction in satellite images, specifically to roads.

4 Results and Discussion

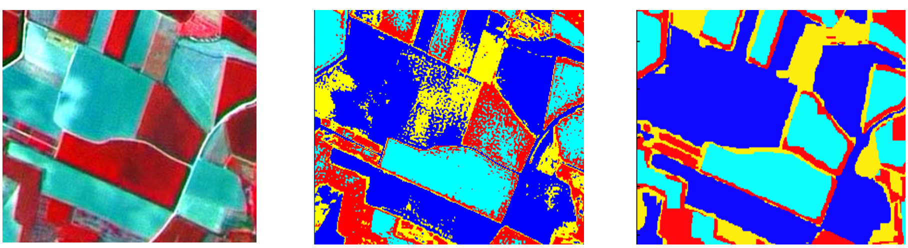
The proposed technique has been applied to imagery obtained by the two high resolution sensors mentioned in the previous section.



(a) (b) (c)

Fig. 3. (a) An original Ikonos 4-band satellite image from the Polytechnic School in Alcala de Henares; (b) An Ikonos image k-means segmentation with 4 classes; (c) Display results obtained with our proposed technique based on graph cut

Fig. 3(a) shows pan sharpening image of the Polytechnic School in Alcala de Henares. It is a false color image (spotlighting vegetation characteristics; a 2, 3, 4 band combination). The k-means algorithm is applied to the 4 bands imagery; the result is shown in Fig. 3(b). Our proposed technique results are presented in Fig. 3(c). Comparing these two classified images, (b) and (c) can be observed as the graph cut algorithm returns a clean classification, minimizing the noise presented on the image segmentation by the k-means, as seen in Fig. 3(b).



(a) (b) (c)

Fig. 4. (a) A 4 bands Spot satellite image over agricultural fields; (b) the k-means (4 classes) cluster results; in (c) results generated by the proposed technique minimizing noise

Fig. 4(a) is false color image (1, 3, 2; see Table 1) from a field area of a SPOT-5 imagery in central Spain; it has also four multispectral bands as the Ikonos. The results of applying the k-mean and the proposed technique are shown in Fig. 4(b) and Fig. 4(c), respectively.

As can be observed, the biggest central field in the image of Fig. 4(a) is classified by the k-means with a lot of noise; however, with the proposed technique, the noise has almost disappeared. This is due to the penalties applied through the constrains by data cost matrix and the smoothness term, which produces a clear separation between regions that represent real geographic units, and to the spatial domain consideration on the graph cut algorithm.

5 Conclusions

A method for segmenting high resolution satellite images based on k-means and graph cuts has been proposed. The method improves the k-means classification by introducing spatial domain information in the energy function of a graph cut algorithm.

The proposed method is useful in extracting cartographic features in thematic cartography; for instance, for land cover and land use classification.

On feature studies, another's functions will be examined on equation (2) applying boundary constrains, and directed graph in order to propose a supervised classification method employing sample areas' treatment.

Acknowledgements

The authors wish to thank to the Spanish National Mapping Agency (Instituto Geográfico Nacional) for provide the images used on this research and to the Spanish MCYT for financial support, project no. CGL2006-07132/BTE. They are also grateful to Bagon Shai for his code and inputs.

References

1. Lillesand, T.M., Kiefer, R.W., Chipman, J.W.: Remote Sensing and Image Interpretation. Wiley, Chichester (2003)
2. Atkinson, P.M.: Resolution manipulation and sub-pixel mapping. In: de Jong, S.M., van der Meer, F.D. (eds.) Remote Sensing Image Analysis: Including the Spatial Domain (Remote Sensing and Digital Image Processing), pp. 51–70. Springer, Heidelberg (2004)
3. Berberoglu, S., Curran, P.J.: Merging Spectral and Textural Information for Classifying Remotely Sensed Images. In: de Jong, S.M., van der Meer, F.D. (eds.) Remote Sensing Image Analysis: Including the Spatial Domain (Remote Sensing and Digital Image Processing), pp. 113–135. Springer, Heidelberg (2004)
4. van der Werff, H., Lucieer, A.: A contextual algorithm for detection of mineral alteration Halos with hyperspectral remote sensing. In: de Jong, S.M., van der Meer, F.D. (eds.) Remote Sensing Image Analysis: Including the Spatial Domain (Remote Sensing and Digital Image Processing), pp. 201–210. Springer, Heidelberg (2004)
5. Carvalbo, L.M., Acerbi, F.W., Clevers Jr., J., Fonseca, L., de Jong, S.M.: Multiscale feature extraction from images using wavelets. In: Steven, M., de Jong, S.M., van der Meer, F.D. (eds.) Remote Sensing Image Analysis: Including the Spatial Domain (Remote Sensing and Digital Image Processing), pp. 237–269. Springer, Heidelberg (2004)

6. Boykov, Y., Funka-Lea, G.: Graph Cuts and Efficient N-D Image Segmentation. International Journal of Computer Vision 70(2), 109–131 (2006)
7. Boykov, Y., Kolmogorov, V.: An Experimental Comparison of Min-Cut/Max-Flow Algorithmsfor Energy Minimization in Vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(9), 1124–1137 (2004)
8. Kolmogorov, K., Zabih, Z.: What Energy Functions can be Minimized via Graph Cuts? IEEE Transactions on Pattern Analysis and Machine Intellegence 26(2), 147–159 (2004)
9. Elias, P., Feinstein, A., Shannon, C.E.: A note on the maximum flow through a network. IRS Transactions on Information Theory 2, 117–119 (1956)
10. Ford Jr., L.R., Fulkerson, D.R.: Flows in Networks. Princeton University Press, Princeton (1962)
11. Greig, D., Porteous, B., Seheult, A.: Exact Maximum A Posteriori Estimation for Binary Images, J. Royal Statistical Soc., Series B 51(2), 271–279 (1989)
12. Lin, M.H.: Surfaces with Occlusions from Layered Stereo. Int'l J. Computer Vision 1(2), 1–15 (1999)
13. Kolmogorov, V., Zabih, R.: Multi-Camera Scene Reconstruction via Graph Cuts. In: Proc. European Conf. Computer Vision, vol. 3, pp. 82–96 (2002)
14. Boykov, Y., Jolly, M.-P.: Interactive Organ Segmentation Using Graph Cuts. In: Proc. Medical Image Computing and Computer-Assisted Intervention, pp. 276–286 (2000)
15. Goldberg, A.V., Tarjan, R.E.: A new approach to the maximum-flow problem. Journal of the ACM (JACM) 35(4), 921–940 (1988)
16. Boykov, Y., Veksler, O., Zabih, R.: Markov random fields with efficient approximations. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 648–655 (1998)
17. The Weizmann Institute of Science Faculty of Mathematics and Computer Science Computer Vision Lab, `http://www.wisdom.weizmann.ac.il/~bagon/matlab.html`
18. Richard, D.O., Hart, E., Peter y Stark David, G.: Pattern Classification. Wiley-Interscience, Hoboken (2004)
19. del-Toro-Almenares, A., Mihai, C., Vanhamel, I., Sahli, H.: Graph Cuts Approach to MRF Based Linear Feature Extraction in Satellite Images. In: Rueda, L., Mery, D., Kittler, J. (eds.) CIARP 2007. LNCS, vol. 4756, pp. 162–171. Springer, Heidelberg (2007)