Anais XIV Simpósio Brasileiro de Sensoriamento Remoto, Natal, Brasil, 25-30 abril 2009, INPE, p. 7909-7916.

Using relevance feedback for classifying remote sensing images

Jefersson Alex dos Santos, Rubens Lampareli, Ricardo da Silva Torres

[1] Institute of Computing – University of Campinas (UNICAMP) – Campinas, SP
[2] Center for Research in Agriculture – UNICAMP – Campinas, SP
jefersson@lis.ic.unicamp.br,rubens@cpa.unicamp.br,rtorres@ic.unicamp.br

Abstract. This paper presents an interactive technique for remote sensing image classification. In our proposal, users are able to interact with the classification system, indicating regions which are of interest. Furthermore, a genetic programming approach is used to learn user preferences and combine image region descriptors that encode spectral and texture properties. The approach to classify images can be divided into four main steps: (i) image partition and region feature extraction, (ii) identification of the partitions which are of interest, (iii) image segmentation, and (iv) region vectorization. This work describes the obtained results from the first two main steps: *partition/extraction of image features* and *recognition of partitions of interest*. So, in the first step the image are partitioned into tiles. Each tile is considered as an independent image and this process starts by the indication of a query image by the user. This query image is assumed to present the same texture and spectral properties of the RSI regions which are of interest. A similarity search is performed and the most similar tiles are returned to the user. The user then indicates if the returned tiles are relevant or non-relevant. By using this feedback, the classification system learns the user needs and tunes itself in order to improve the results in the next iteration. This process is repeated until the user is satisfied with the result. Experiments demonstrate that the proposed method is effective and suitable for image classification tasks.

Keywords: remote sensing image classification, relevance feedback, genetic programming ..

1. Introduction

Brazilian agriculture has obtained efficient, competitive, and dynamic results. In the last decade, agriculture has increased its contribution to the Brazilian Gross Domestic Product (GDP), representing around 10% of the total GDP. In this scenario, there is a huge demand for information systems to support monitoring and planning of agriculture activities in Brazil. One of the most used approaches for crop monitoring is based on the use of Remote Sensing Images (RSIs).

RSIs provides the basis for the creation of information systems that support the decision-making process based on soil occupation changes. In these systems, two important issues need to be addressed: how to identify (recognize) regions of interest and, later, how to extract/define polygons around these regions.

The identification and polygon extraction tasks usually rely on applying classification strategies that exploit visual aspects related to spectral and texture patterns identified in RSI regions. These tasks can be done automatically or manually.

The “manual" approach is based on image editors where users can define or draw polygons that represent regions of interest using the raster image as background. The extraction of polygons from raster images is called *vectorization*.

In general, automatic approaches use classication strategies based on pixel information (SHOWENGERDT, 1983). The main drawback of these approaches is concerned with its sensitivity to image noises (e.g., for example, distortions that can be found in mountainous regions). Another important problem in the automatic approaches is concerned with the fact that they usually fail to correct identify borders between distinct regions within the same image. Thus, in practical situations, the results obtained need to be revised. As these revisions take a lot of time, it is sometines more convenient to the user to perform recognition manually.

This paper addresses these shortcomings by presenting a semi-automatic approach for RSI classification. The proposed solution relies on the use of an interactive strategy, called *relevance feedback* (ZHOU; HUANG, 2003), based on which the classification system can learn which regions are of interest. The proposed image classification process with relevance feedback is comprised of four steps: (i) showing a small number of retrieved image regions to the user; (ii) user indication of relevant and non-relevant regions; (iii) learning the user needs from her feedbacks; (iv) and selecting a new set of regions to be shown. This procedure is repeated until a satisfactory result is reached.

In this paper a recently proposed relevance feedback method for interactive image search (FERREIRA et al., 2008; SANTOS; FERREIRA; TORRES, 2008) is used for image classification. This method adopts a genetic programming approach to learn user preferences in a query session. Genetic programming (GP) (KOZA, 1992) is a Machine Learning technique used in many applications, such as data mining, signal processing, and regression (BHANU; LIN, 2004; FAN; GORDON; PATHAK, 2004; ZHANG et al., 2004). This technique is based on the evolution theory and aims to find near optimal solutions. The use of GP is motivated in this work by the previous success of using this technique in information retrieval (FAN; GORDON; PATHAK, 2004) and Content-Based Image Retrieval (CBIR) (TORRES et al., 2008) tasks.

2. Related Work

2.1. Classification of RSIs

Images provided by satellite sensors have been used in large scale for crop monitoring and production predictions. However, there is not a satisfactory method to classify RSIs so far. Terrain distortions and the interference of clouds for example, make classification a hard problem. Another issue is to provide effective classification strategies considering the different evolution stages of a crop. Traditional classification methods are based on pixel analysis. The most used pixel classification algorithm, MaxVer (SHOWENGERDT, 1983), however, is not very effective. Several new methods have been proposed to improve the performance of MaxVer-based techniques. In (MO et al., 2007), a new method considering image segmentation, GIS, and data mining algorithms was presented. Compared with pixel-based classification, the results showed best agreement with visual interpretation. The work proposed in (YILDIRIM; ERSOY; YAZGAN, 2005) applied a morphological filter in an image which was classified by MaxVer algorithm. The results were compared with the other classification algorithms (Fisher linear likelihood, minimum Euclidean distance and ECHO). In (KIM; JEONG; PARK, 2007), three Land Cover Classification Algorithms are compared for monitoring North Korea using multi-temporal data.

2.2. CBIR and Relevance Feedback

CBIR systems provide efficient and effective means to retrieve images. In these systems, the searching process consists in, for a given image, computing the most similar images stored in the database. The searching process relies on the use of image *descriptors*. A descriptor can be characterized by two functions: *feature vector extraction* and *similarity computation*. The feature vectors encode image properties, like color, texture, and shape. Therefore, the similarity between two images is computed as a function of their feature vectors distance.

In some CBIR approaches the descriptors are statically combined, that is, the descriptors composition is fixed and used in all retrieval sessions. Nevertheless, different people can have distinct visual perception of a same image. Motivated by this limitation, *relevance feedback* approaches were incorporated into CBIR systems (RUI et al., 1998; COX et al., 2000; HONG; TIAN; HUANG, 2000). This technique makes possible the user interaction with the retrieval systems.

3. The Semi-automatic Vectorization Approach

The proposed vectorization approach can be divided into four main steps: (i) image partition and region feature extraction, (ii) identification of the partitions which are of interest, (iii) image segmentation, and (iv) region vectorization. Figure 1 ilustrates the steps of the vectorization process.

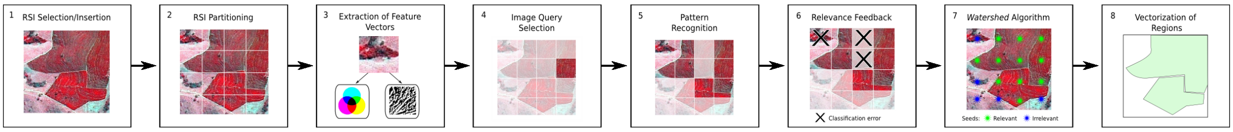


Figure 1: Steps of the proposed vectorization process.

Let I be an RSI and $I_{i\times j}$ a sub image of I composed by $n \times n$ pixels. The *image partition process* consists of creating a grid of $n \times n$ sub images (tiles) from I. The value of n is based on the estimated size of a region of interest. This way, an ideal value to n is that one which makes the sub images be found inside regions of interest. For each subimage of I, spectral and texture features are extracted using pre-defined image descriptors.

The process of *identifying relevant partitions* exploits a relevance feedback strategy. Each tile is considered as an independent image and this process starts by the indication of a query image by the user. This query image is assumed to present the same texture and spectral properties of the RSI regions which are of interest. A similarity search is performed and the most similar tiles are returned to the user. The user then indicates if the returned tiles are relevant or non-relevant. By using this feedback, the classification system learns the user needs and tunes itself in order to improve the results in the next iteration. This process is repeated until the user is satisfied with the result.

In our proposal, genetic programming is used in the learning process. The method used in this paper is based on the proposals described in (FERREIRA et al., 2008; SANTOS; FERREIRA; TORRES, 2008). In (FERREIRA et al., 2008), Ferreira et al. proposed a GP-based approach for relevance feedback in CBIR systems. In (SANTOS; FERREIRA; TORRES, 2008), this approach is extended to handle image region descriptors.

After the tiles of interest are identified, the next step is concerned with the *segmentation* of relevant regions. The segmentation process of the image is performed by using a "watershed"-based (LOTUFO; FALCÃO, 2000) algorithm. This algorithm segmentates images using seeds. The seeds are based on areas of interest identified in the last step.

Finally, the vectorization process consists of using the segmented image to extract the polygons of the region of interest.

This work describes the obtained results from the first two main steps: *partition/extraction of image features* and *recognition of regions of interest*. Details about the process of classification process are described in in (SANTOS; FERREIRA; TORRES, 2008).

4. Validation

This section describes the experiments performed to validate our method to identify and classify regions.

4.1. Remote Sensing Images

Two RSIs are used to validade our method. One can be classified such as "easy recognition" (pasture image) while the other is "hard recognition" (coffee). Information about used RSIs is showed in Table 1.

Table 1: Remote Sensing Images used in the experiments.

	Image1	**Image2**
Recognition level	easy	hard
Region of interest	pasture	coffee
Terrain	plain	mountainous
Satelite	CBERS	SPOT
Spatial resolution	20 meters	2,5 meters
Bands composition	R-IR-G (342)	IR-NIR-R (342)
Acquisition date	08–20–2005	08–29–2005
Location	“Laranja Azeda" Basin, MS	Monte Santo County, MG
Dimensions (px)	1310 × 1842	2400 × 2400

Table 2: Image descriptors used in the experiments.

Descriptor	**Type**
Color Histogram (SWAIN; BALLARD, 1991)	Color
Color Moments (STRICKER; ORENGO, 1995)	Color
BIC (STEHLING; NASCIMENTO; FALCÃO, 2002)	Color
Gabor Wavelets (LEE, 1996)	Texture
Spline Wavelets (UNSER; ALDROUBI; EDEN, 1993)	Texture

4.2. Image Descriptors

To validate our method, texture and spectral properties from the images are combined. To extract information about spectral properties from the images, color-based descriptors were used. Texture descriptors were used to extract texture properties from the images. Table 2 shows the used descriptors.

4.3. Baselines

We compare our method against *Maximum Likelihood (MaxVer) Classification* (SHOWENGERDT, 1983). It is the most common supervised classification method used with remote sensing image data. It is considered as a parametric algorithm and it assumes a particular class statistical distribution, commonly the normal distribution. The implementation of MaxVer algorithm requires the computation of the probability that each pixel belongs to each of the defined classes. Each pixel is then assigned to that class for which that probability is the greatest.

4.4. Implementation

The system is implemented with the minimal requirements to validate our method. The recognition of partitions of interest requires the definition of several GP parameters (e.g., mutation rate, population size). We used the same parameters as reported in (SANTOS; FERREIRA; TORRES, 2008).

4.5. Effectiveness Measure

As aforementioned, the classification results of the proposed method are related with the number of user interactions. Thus, to improve our method, we use *kappa–interactions* curves. The kappa is an effective index to compare classified images. To calculate it is necessary to create an *error matrix*. An *error matrix* is a square array of numbers set out in rows and columns. It expresses the number of sample units (pixels, clusters, or polygons) assigned to a particular category in one classification relative to the number of sample units assigned to a particular category in another classification (CONGALTON; GREEN, 1977). The matrix error is a very effective way to represent map accuracy in that the individual accuracies of each category are plainly described along with both the errors of inclusion (commissions errors) and errors of exclusion (omission errors) present in the classification. A commission error is simply defined as including an area into a category when it does not belong to that category. An omission error is excluding that area from the category in which it truly does belong. The kappa index does not use just the elements from the main diagonal of the error matrix, it includes all the elements.

4.6. Experiment Design

In our experiments we fixed the tile size according to the common extension value of a *region of interest*. Coffee crops are normally in small parcels on the same farm. We defined that 75×75 $meters$ is a good value to the size of the partition. To pasture parcels, that are larger, the chosen value was 400×400 $meters$. The dimension of partitions are fixed in experiments. We used 30×30 $pixels$ to partition the coffee image and 20×20 $pixels$ for the pasture image. The number of partitions for the pasture and coffee images were 5980 and 6400, respectvively.

We used a “mask" contained all regions of interest from the RSIs used in the experiments. A “mask" is a binary image where value 1 represents pixels of regions of interest. The “masks" used in our experiments are classified mannually by agricultural specialists.

The user interaction was simulated. To do it, we created a *groundtruth* based on the “mask". The *groundtruth* is a binary image. A partition is represented by pixels with value 1, if the number of relevant pixels is bigger than a percentual index. In our experiments the index was fixed in 50%. The number of partitions showed to the user on each iteration was 20.

The proposed technique to recognize regions (section (FERREIRA et al., 2008; SANTOS; FERREIRA; TORRES, 2008)) creates a ranking of partitions based on their similarity with regard the reference image defined by the user. On the other hand, the results have to be a binary image representing relevants and irrelevants partitions. Thus, we made experiments using different *thresholds* to separate relevant and irrelevant partitions: 5%, $7,5\%$, 10%, 15%, 20%, and 30% of the top ranked partitions. To the experiment recognizing coffee, we tested also the thhreshold value 40%.

As aforementioned in section 4.3, the proposed method is compared with *MaxVer*. The Image 1 was classified by *MaxVer* with probability threshold 0.8 and using 20.580 points of pasture sample. The Image 2 was classified with probability threshold 0.98 and using 43.630 points of coffee sample.

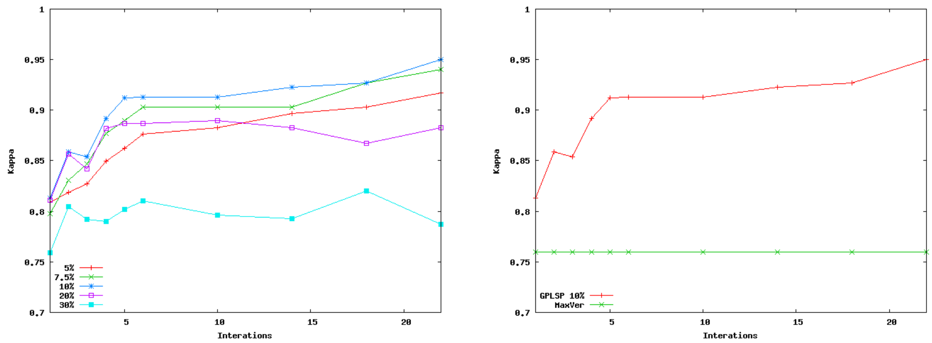
4.7. Results

Figure 2(a) refers to the pasture image. It shows curves related to the kappa index variation along iterations, considering different threshold applied to the ranked partitions: 5%, 7.5%, 10% , 20%, and 30%. Figure 2(b) shows the best curve in the Figure 2(a) and the kappa index obtained using the *MaxVer* image classification.

Figure 3 ilustrates the original RSI, the mask (ground truth) and the classifications using considering different threshold values to pasture recognition.

Figure 4 (a) shows the kappa-iterations curves for the coffee image. Figure 4(b) shows the best line curve in the Figure 4 (a) and the kappa index obtained using the *MaxVer* image classification.

Figure 5 ilustrates the original RSI, the mask (ground truth) and the classifications using considering different threshold values to coffee recognition.



(a) (b)

Figure 2: Kappa-Iterations curves considering the pasture image. (a) The proposed classification method considering 5%, 7.5%, 10%, 20%, and 30% threshold values. (b) The best curve showed in (a) – considering a threshold value of 7.5% – and the $MaxVer$ classification accuracy.

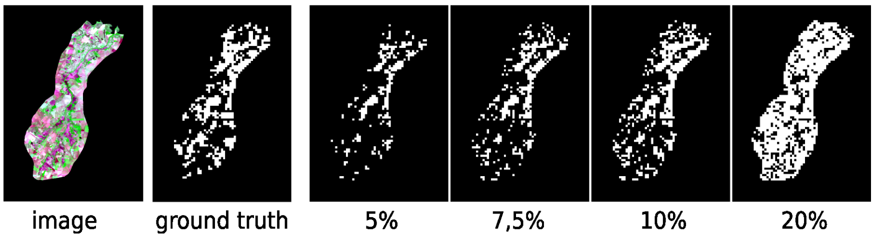
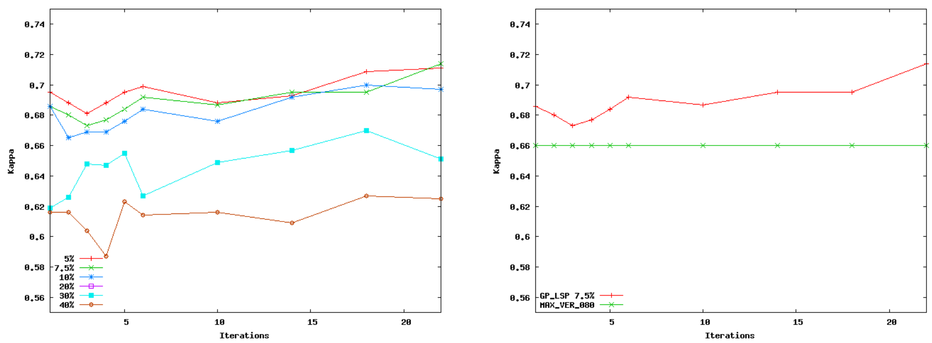


Figure 3: Steps of the proposed vectorization process.



(a) (b)

Figure 4: Kappa-Iterations curves considering the coffee image. (a) The proposed classification method considering 5%, 7.5%, 10%, 20%, and 30% threshold values. (b) The best curve showed in (a) – considering a threshold value of 7.5% – and the $MaxVer$ classification accuracy.

5. Conclusions

We have presented a relevance feedback approach to classifiy remote sensing images. The proposed method uses genetic programming to learn the user needs by taking into account her feedback and spectral and texture properties of regions which are encoded by image descriptors Experiments showed that the proposed method is suitable to recognize regions of interest, presenting better accuracy than the traditional MaxVer method. The next stage of our work is to implement a watershed-based segmentation algorithm and compare the results with other classification approaches. We also plan to make new experiments using different image

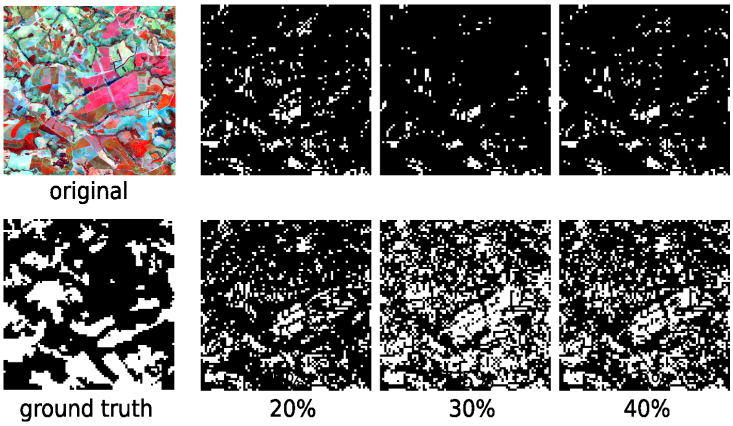


Figure 5: Steps of the proposed vectorization process.

collections.

6. Acknowledgments

References

BHANU, B.; LIN, Y. Object Detection in Multi-Modal Images Using Genetic Programming. *Applied Soft Computing*, v. 4, n. 2, p. 175–201, May 2004.

CONGALTON, R. G.; GREEN, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices.* Washington, DC: Lewis Publishersr, 1977.

COX, I. J. et al. The Bayesian Image Retrieval System, PicHunter: Theory, Implementation, and Psychophysical Experiments. *IEEE Transactions on Image Processing*, v. 9, n. 1, p. 20–37, January 2000.

FAN, W.; GORDON, M. D.; PATHAK, P. A generic ranking function discovery framework by genetic programming for information retrieval. *Information Processing & Management*, v. 40, n. 4, p. 587–602, July 2004.

FERREIRA, C. et al. Image retrieval with relevance feedback based on genetic programming. In: *23nd Brazilian Symposium on Database*. Campinas, SP, Brasil: [s.n.], 2008.

HONG, P.; TIAN, Q.; HUANG, T. S. Incorporate support vector machines to content-based image retrieval with relevant feedback. In: *Proceedings of the 7th IEEE International Conference on Image Processing*. [S.l.: s.n.], 2000. p. 750–753.

KIM, D. hyung; JEONG, S. gyu; PARK, C. hwa. 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*, 2007.

KOZA, J. R. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA, USA: MIT Press, 1992.

LEE, T. S. Image representation using 2d gabor wavelets. *IEEE Transactions Pattern Analysis Machine Intelligence*, v. 18, n. 10, p. 959–971, 1996. ISSN 0162-8828.

LOTUFO, R.; FALCÃO, A. The ordered queue and the optimality of the watershed approaches. In: *In Mathematical Morphology and its Applications to Image and Signal Processing*. [S.l.]: Kluwer Academic Publishers, 2000. p. 341–350.

MO, D.-K. et al. Design and implementation of a high spatial resolution remote sensing image intelligent interpretation system. *Data Science Journal*, v. 6, p. S445–S452, 2007.

RUI, Y. et al. Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 8, n. 5, p. 644–655, 1998.

SANTOS, J. A. dos; FERREIRA, C.; TORRES, R. A genetic programming approach for relevance feedback in region-based image retrieval systems. In: *Brazilian Symposium on Computer Graphics and Image Processing, 21 (SIBGRAPI)*. Campo Grande, MS, Brasil: [s.n.], 2008.

SHOWENGERDT, R. *Techniques for Image Processing and Classification in Remote Sensing*. New York: Academic Press, 1983.

STEHLING, R.; NASCIMENTO, M.; FALCÃO, A. A Compact and Efficient Image Retrieval Approach Based on Border/Interior Pixel Classification. In: *Proceedings of the 11th ACM International. Conference on Information and Knowledge Management*. McLean, Virginia, USA: [s.n.], 2002. p. 102–109.

STRICKER, M. A.; ORENGO, M. Similarity of Color Images. In: *Storage and Retrieval for Image and Video Databases (SPIE)*. [S.l.: s.n.], 1995. p. 381–392.

SWAIN, M.; BALLARD, D. Color Indexing. *International Journal of Computer Vision*, v. 7, n. 1, p. 11–32, 1991.

TORRES, R. da S. et al. A genetic programming framework for content-based image retrieval. *Pattern Recognition*, 2008.

UNSER, M.; ALDROUBI, A.; EDEN, M. A family of polynomial spline wavelet transforms. *Signal Process.*, v. 30, n. 2, p. 141–162, 1993.

YILDIRIM, I.; ERSOY, O. K.; YAZGAN, B. Improvement of classification accuracy in remote sensing using morphological filter. *Advances in Space Research*, 2005.

ZHANG, B. et al. Combining structural and citation-based evidence for text classification. In: *Proceedings of the 13th ACM Conference on Information and Knowledge Management*. [S.l.: s.n.], 2004. p. 162–163.

ZHOU, X. S.; HUANG, T. S. Relevance feedback in image retrieval: A comprehensive review. *Multimedia Systems*, v. 8, n. 6, p. 536–544, 2003.