

# Toward the Labeled Segmentation of Natural Images Using Rough-Set Rules

Fernando J. Navarro-Avila, Jonathan Cepeda-Negrete,  
and Raul E. Sanchez-Yanez<sup>(✉)</sup>

Universidad de Guanajuato DICIS, Salamanca, Guanajuato, Mexico  
{jf.navarroavila,j.cepedanegrete,sanchezy}@ugto.mx

**Abstract.** This article introduces an approach that integrates color and texture features for the segmentation of natural images. In order to deal with the vague or imprecise information that is typically shown in this kind of scenes, our method consists in a supervised classifier based on rules obtained using the rough-set theory. Such rough classifier yields a label per pixel using as inputs only three color and three textural features computed separately. These labels are used to carry out the image segmentation. When comparing quantitatively the results from this work with state-of-the-art algorithms, it has shown to be a competitive approach to the image segmentation task. Moreover, the labeling of each pixel offers advantages over other segmentation algorithms because the outcome is intuitive to humans in two senses. On one hand, the use of simple rules and few features facilitate the understanding of the segmentation process. On the other hand, the labels in the segmented outcomes provide insight into the image content.

**Keywords:** Rough-set rule · Supervised classifier · Labeled segmentation · Natural image

## 1 Introduction

Image segmentation is defined as the process of partitioning an image into its constituent components, called segments or clusters of pixels [1]. The resulting segments are collections of pixels that share a similarity with respect to some feature or property. Over the last years, image segmentation has resulted to be of the most difficult tasks in computer vision and image processing, but also a promising approach towards object recognition and image understanding, mainly because each region within an image usually corresponds to an object. According to Ilea and Whelan [2], the image segmentation task has been developed mostly using color and texture properties in a large number of approaches, including region growing, edge detection, clustering and histogram-based, among others. In fact, the current research is leaning toward the integration of features and methods.

The use of color and texture information collectively, has strong links with the human perception and, in many scenarios, the use of color or texture solely

is not sufficient to describe the image content. The segmentation of natural images exemplifies this problem, mainly because these images present significant inhomogeneities in color and texture [2, 3]. In addition, scenes are often complex, with a degree of randomness and irregularity. Likewise, the strength of texture and color attributes can vary considerably from image to image when adding distortions due to uneven illumination, scale changes, and other sources.

Rough set theory, proposed by Pawlak in [4], is a promissory approach to deal with inconsistency of data. The theory has been proposed to solve a wide number of problems, particularly those related to artificial intelligence and cognitive science as machine learning, data mining and pattern recognition, but applications in image processing are relatively few. The rough set theory offers advantages when compared to other approaches. (I) There is no need of additional or preliminary information regarding the data. (II) Data quantification is carried out together with the rule induction. (III) Redundant information is reduced in two ways: when identical objects are represented several times or when some attributes do not contribute to the classification (superfluous attributes). (IV) The minimum set of attributes that preserves the knowledge in the original data is found. (V) Decision rules are induced automatically from data. (VI) The rough set theory can ease the interpretation of results.

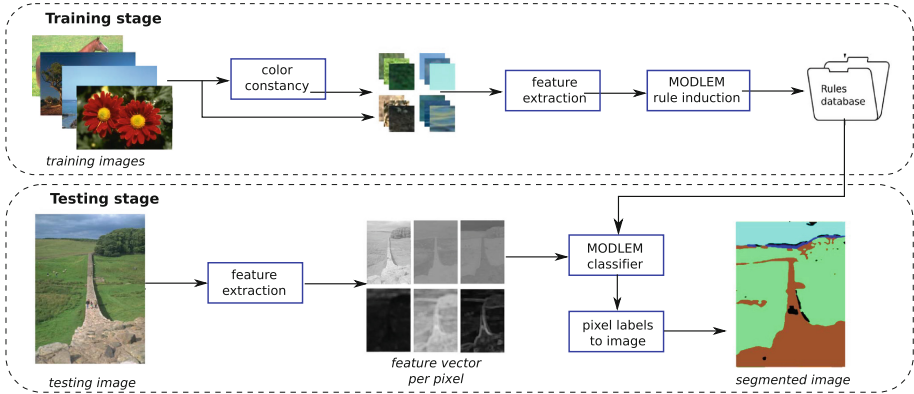
Although several methods of color-texture image segmentation have been proposed, rule-based classifiers for image segmentation have not been widely explored. Our approach takes into account the fact that natural images are taken under different light conditions and other irregularities present by the nature of image acquisition. At first, we use color constancy in an image preprocessing step, and afterward a rough-set-rule classifier is utilized as an alternative to image segmentation approaches. This classifier integrates features allowing the processing of imprecise, incomplete or uncertain information of the real world. Furthermore, a rule-based classifier induces intuitive decision rules, thus enabling the understanding of the image segmentation process. Besides, the labeling of pixels, and consequently, of regions, provides a higher level of knowledge regarding the image content.

The rest of this paper is organized as follows. Section 2 describes the methodology used throughout our work. It discusses the selection of color and texture features, and the rough set rules obtained by the classifier. Section 3 includes the experimental results in the test series, and the observations from the data obtained. Finally, the concluding remarks are given in Sect. 4.

## 2 Methodology

In this section, we describe the methodology implemented for the natural image segmentation task. An explanation of the process is provided, as well as the criteria to choose the methods and the parameters used for the system configuration.

The block diagram shown in Fig. 1 depicts the system used for image segmentation. Regarding this diagram, two general processes are distinguished: training and testing stages.



**Fig. 1.** Block diagram of the rough-based system for image segmentation.

The purpose of the training stage is to induce decision rules in order to generalize and label future unknown pixels using color and texture features. The set of training images is constituted by instances to be described. Each input image is a representative sample of a desired class. In our study, the number of classes is limited to four (“Foliage”, “Soil”, “Sky” and “Water”) and, additionally, an “Indeterminate” class is included for denoting outliers. At the training stage, the first step consists in the application of a color constancy algorithm to the input image. In this study, the Gray-World assumption [5] and the White-Patch [6] algorithms are used to process the training images and increase the number of samples.

## 2.1 Feature Extraction

The second step consists of the feature extraction. The extraction of color features in this work requires a color space that shows minimum correlation between components, as a consequence providing more information of the scene and thus improving the segmentation. A color space with perceptual characteristics in its components is needed. In other words, little changes in each component are translated into changes perceived by humans. The CIELAB color space has the proper characteristics for this work. This color space considers one channel for Luminance  $L$  and two color channels  $a$  and  $b$ . We use the well-known equations in literature and the D65 white reference [7]. The image is transformed from the RGB color space to CIELAB giving as a result three images, each containing one of the color features used in the system ( $L$ ,  $a^*$ ,  $b^*$ ). These three images are smoothed using a simple mean filter.

In addition to color features, the textural content exhibited in all images plays an important role in image analysis in a wide range of applications as medical imaging, industrial processes and robotics. Despite texture is an intuitive concept, there is no formal definition for a visual texture. Tamura *et al.* [8] mention that “A texture can be considered as a macroscopic region. Its structure is attributed to repetitive or primitive patterns whose elements are arranged according to a rule of position.”

The processed image after the color constancy algorithm is transformed to gray scale, and the texture features are extracted on a neighborhood centered over each pixel using a  $w \times w$  window. Homogeneity and entropy are extracted using sum and difference histograms [9] and the standard deviation map is extracted as well, obtaining three texture maps. These images are also smoothed using a low-pass filter. This procedure is repeated for each image in the training set obtaining an  $m \times n$  matrix, where  $m$  is the number of samples and  $n$  is the number of features. In our study,  $n$  equals to six, combining only three color features and three texture features.

For computing the Sum and Difference Histograms (SDH) [9], the relative displacement vector ( $\mathbf{V}$ ) between two picture elements is an SDH parameter, and in our study, it is defined as the composition of the Cartesian product  $R \times \theta$ , where  $R = \{1, 2\}$  and  $\theta = \{0, \pi/4, \pi/2, 3\pi/4\}$ . Homogeneity and entropy are calculated using

$$\text{Entropy} = - \sum_i P_s(i) \cdot \log(P_s(i)) - \sum_j P_d(j) \cdot \log(P_d(j)) \quad (1)$$

$$\text{Homogeneity} = \sum_j (1 + j^2)^{-1} \cdot P_d(j) \quad (2)$$

where  $P_s$  and  $P_d$  are the normalized SDH.

The last texture feature used throughout the system is a standard deviation map. Texture regions within an image show more intensity variations than those pixels in homogeneous regions. A measure of those variations is used to determine the boundaries of such textured regions. Moreover, different textures have distinct variations in intensity allowing to discriminate those textures.

The computation of the standard deviation image  $T$  is obtained for each pixel in the image in a desired neighborhood consisting of a square region containing  $k = (2d + 1)^2$  pixels and centered at the current pixel position. The parameter  $d$  is the number of pixels from the central pixel to a side of the window. Equations (3) and (4) are used to compute  $\sigma$ . In these equations  $I_i$  represents the intensity of the  $i_{th}$  pixel of the neighborhood,  $\mu$  and  $\sigma$  are the first and second statistical moments about zero, respectively.

$$\mu = \frac{1}{k} \sum_{i=1}^k I_i, \quad (3)$$

$$\sigma = \sqrt{\sum_{i=1}^k (I_i - \mu)^2 / (N - 1)}. \quad (4)$$

Now that the features have been defined, the system requires a pixel wise classifier in order to assign labels to pixels according to the input features. In the training stage, the last step consists of the induction of rules using MODLEM [10], which is an algorithm based on rough set theory. The rules induced in this step are stored on a database of rules.

The testing stage is similar to the training one. For each testing image, the feature extraction is performed using the same window size, but this time the resulting matrix is of the size  $P \times N$ , where  $P$  is equal to the number of pixels within the test image and  $N$  still being the number of features. Each pixel generates a feature vector. The matching of this vector with the stored rules yields a label. Therefore, the process of segmentation involves the labeling of pixels at each position on the testing image, obtaining a matrix of labels with the same size that the image. The last step in the testing stage is the conformation of a map of labels as an image.

## 2.2 The MODLEM Classifier

After the feature extraction process, a decision table is obtained. This decision table can be inconsistent due to unbalanced data or confusing samples. Therefore, decision rules can be generated from rough approximations [11]. For this matter, rule-based classifiers are suitable algorithms. These algorithms iteratively create a set of rules for each class or concept. In addition, one of the main advantages of using rule-based classifiers is that a comprehensive description for the class or the concept is generated.

The rules allow the control of the complexity of the classifier describing the class, in order to simplify the understanding of the classification process, making this procedure intuitive to humans. Usually, decision rules are of the form IF (*antecedent*) THEN (*consequence*). The left-hand side of the rule is called antecedent; this expression refers to an attribute and its value. The right-hand side of the rule defines the consequence, in our case the label given for the class.

In this work, the MODLEM classifier is used. It is a sequential covering algorithm introduced by Stefanowski in [12] as a modification of the induction rule algorithm developed by Grzymala-Busse [13]. This rule induction algorithm has been chosen because there is no need for data discretization as it is computed simultaneously with the rule induction.

This algorithm generates a minimal set of decision rules for every decision concept (decision class or its rough approximation in case of inconsistent examples). Such a minimal set of rules attempts to cover all positive examples of the given decision concept denoted as  $B$ , and not to cover any negative examples ( $U \setminus B$ ).

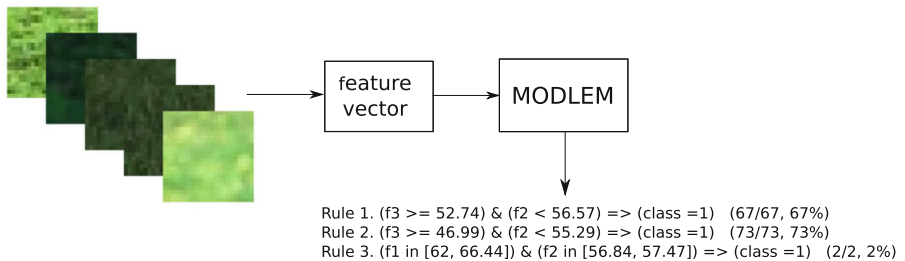
The main rule induction scheme is described in the following steps:

1. Create a first rule by choosing sequentially the “best” elementary conditions according to the chosen “*Find best condition function*”.
2. When the rule is stored, all learning positive examples that match this rule are removed from consideration.
3. The process is repeated while some positive examples of the decision concept remain still uncovered.
4. Then, the procedure is sequentially repeated for each set of examples from a concept or category.

After the rules are induced in the training stage, then the testing stage consists of using those rules to predict a class assignment for pixels in an unseen

image. This is done evaluating the matching of the new feature description with the condition parts of decision rules. This may result in unique matching to rules from one single class. However, two other ambiguous cases are possible: matching to more rules indicating different classes or the feature description does not match any of the rules. In these cases, it is necessary to apply proper strategies to solve these conflict cases [14].

The rule induction process is illustrated in Fig. 2, completing the whole system. This figure describes the classifier training stage, where the input to the classifier is a joint color-texture vector feature per sample to be classified. In the figure, the rule induction for the “Foliage” class is shown.



**Fig. 2.** Sample of rule induction using MODLEM (Color figure online)

To measure our system performance, an evaluation strategy is needed. This evaluation consists of a standard dataset, an evaluation measure and a specific strategy for training and testing stages. This strategy is discussed in the following section.

### 3 Experimental Results

In this section, we present the results obtained by our method called Labeled Rough-Set-Based Segmentation (LRS). The goal of test series has been to evaluate the performance of our approach in comparison with other state-of-the-art algorithms. In these tests, we have applied our approach to natural images using four categories: “Foliage”, “Soil”, “Sky” and “Water”.

For the experiments, the Berkeley segmentation dataset (BSDS300) [15] is used. This dataset consists of 300 natural images, either color or grayscale. As our method utilizes color and texture features, only the color images have been used. Moreover, for each image in the dataset, a set of ground-truth images made by humans is available, and it is used to quantify the reliability of a given method. The dataset also presents a diversity of content, including landscapes and portraits, resulting in a challenge for any segmentation algorithm.

In order to assess the segmentation performance of our method, three widely used metrics are adopted: Probabilistic Rand Index (PRI) [1], Variation of information (VoI) [16] and Global Consistency Error (GCE) [15]. Each metric provides a measure obtained from the comparison of a segmented image and the corresponding ground-truth. We have used the source code provided by Yang *et al.* [17] to evaluate our system performance using PRI, VoI and GCE.

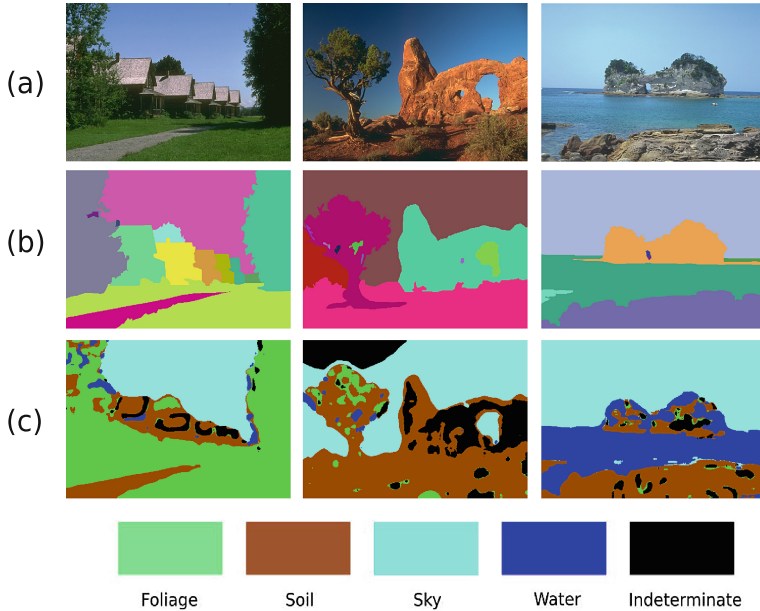
As a general explanation, the PRI counts the number of pixel pairs whose pixels are consistent, for both, the ground-truth and the segmentation result. If the segmented image and the ground-truth images have no matches, then the PRI is zero, giving the minimum value. On the other hand, the maximum value of 1 is achieved if both, the outcome and all the ground-truth images are identical. According to Yang *et al.* [17], the PRI describes a good correlation with the human perception through the hand-labeled segmentations. The VoI metric calculates conditional entropies between distributions of class labels. The GCE evaluates the extent to which a segmentation map can be considered as a refinement of another segmentation. For both, the VoI and the GCE metrics, the segmentation is better if the values are closer to zero.

### 3.1 Qualitative Evaluation

Three parameters are of special interest in our study: (i) the inspection window to compute the texture features (homogeneity and entropy); (ii) the inspection window for the standard deviation map and (iii) the number of samples used to train the pixel classifier. We have exhaustively searched for these parameters, finding to be the best a  $23 \times 23$  window for the first parameter and an  $11 \times 11$  window for the second one. Regarding the number of samples, it has been found that using more than 90 representative samples per class does not improve the segmentation results. It is important to mention that a post-processing stage is performed after the segmentation results are obtained. It consists of a simple median filter of  $3 \times 3$  size applied to the segmented image. The final goal of this procedure is to eliminate extremely small regions and improve segmentation results.

The segmentation approach proposed in this work has been carried out with the 300 images from the BSDS300. Three examples are given in Fig. 3 where the images in the row (a) correspond to original images taken from the Berkeley dataset. The row (b) shows human-made segmentations taken from the ground-truth images. Notice that the reference images were colored randomly. The row (c) shows the images obtained by our system. Here, each pixel is associated to a class. Green, brown, light blue and dark blue correspond to “Foliage”, “Soil”, “Sky” and “Water” classes, respectively. Those pixels that the system is unable to classify based on the given features are colored in black.

It is important to point out some details of our results. The outcome obtained in row (c) of Fig. 3 shows a segmentation similar to the ground-truth. Our system is able to classify the pixels within the image based on color and texture features in a correct manner. However, our system cannot distinguish grass from trees based on the features used. In the same manner, the ability of abstraction of



**Fig. 3.** Example showing qualitative results: (a) Input image, (b) Human-made reference segmentation (not labeled) and (c) Segmented image using our approach. Last row: Color labels for each class in our approach (Color figure online).

a human to say if a house is in front of another as shown in the ground-truth image is a task that is beyond the scope of this work.

The results show that the final goal of this work has been achieved. Those images containing the trained classes are well segmented. The integration of color and texture using rough-set-rules was a successful approach for the image segmentation task.

### 3.2 Performance Evaluation

In our evaluation, we have compared our approach (LRS) with three state-of-the-art algorithms that use both, color and texture features: the J-image Segmentation (JSEG) method proposed by Deng and Manjunath [18], the Compression-based Texture Merging (CTM) approach introduced by Yang *et al.* [17], and the Clustering-based image Segmentation by Contourlet transform (CSC) presented by An and Pun [19]. The average performance of each method using the three quantitative measures, PRI, VoI and GCE is presented in Table 1.

Under specific circumstances, the results obtained by our method are comparable to those obtained using state-of-the-art methods, as can be seen in Table 1. Considering that these results have been obtained using a reduced number of classes to partition the universe of features, and that the qualitative results shown in Fig. 3 are good, our method is well placed among other approaches.



**Table 1.** Average performance of the segmentation algorithms.

Method	PRI	VoI	GCE
JSEG (300)	0.774	<b>2.134</b>	0.196
CTM $_{\gamma=0.1}$ (300)	0.756	2.464	<b>0.176</b>
CSC (100)	<b>0.796</b>	2.732	0.225
LRS (300)	0.602	2.529	0.254

Besides, for many real-world applications, it can be more important to have labeled clusters of pixels than achieving a higher accuracy in segmentation results.

## 4 Conclusions

The segmentation based on rough-set-rules has resulted to be a promissory tool for the integration of color and texture features. Our method has proven to be an efficient and robust approach for feature integration. Even in those cases that our quantitative results are below the state-of-the-art approaches, our system provides an important feature over other methodologies: for each pixel within the image, there is a related label that corresponds to the pixel categorization. This means that our results not only provide the segmented image, but each segment of the image represents a class, which is intuitive to humans. Having a label for each segment provides basic information to make further assertions about the image, for example, if the segment with label  $X$  is besides, above, below or under another segment with label  $Y$ . We believe that our approach is a first step toward image understanding.

**Acknowledgments.** Jonathan Cepeda-Negrete thanks the Mexican National Council on Science and Technology (CONACyT), scholarship 290747 (Grant No. 388681/254884) and to the University of Guanajuato (PIFI-2015 program) for the financial support provided.

## References

1. Unnikrishnan, R., Pantofaru, C., Hebert, M.: Toward objective evaluation of image segmentation algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(6), 929–944 (2007)
2. Ilea, D.E., Whelan, P.F.: Image segmentation based on the integration of colour-texture descriptors - a review. *Pattern Recognit.* **44**(10), 2479–2501 (2011)
3. Hossein, M., Shankar, R.R., Allen, Y.Y., Shankar, S.S., Yi, M.: Segmentation of natural images by texture and boundary compression. *Int. J. Comput. Vis.* **95**, 86–98 (2011)
4. Pawlak, Z.: Rough sets. *Int. J. Comput. Inf. Sci.* **11**(5), 341–356 (1982)

5. Buchsbaum, G.: A spatial processor model for object colour perception. *J. Frankl. Inst.* **310**(1), 1–26 (1980)
6. Land, E.H., McCann, J.: Lightness and retinex theory. *J. Opt. Soc. Am.* **61**(1), 1–11 (1971)
7. Schanda, J.: *Colorimetry: Understanding the CIE System*. Wiley, Hoboken (2007)
8. Tamura, H., Mori, S., Yamawaki, T.: Textural features corresponding to visual perception. *IEEE Trans. Syst. Man Cybern.* **8**(6), 460–473 (1978)
9. Unser, M.: Sum and difference histograms for texture classification. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**(1), 118–125 (1986). (PAMI)
10. Stefanowski, J.: The rough set based rule induction technique for classification problems. In: *Proceedings of 6th European Conference on Intelligent Techniques and Soft Computing EUFIT*, vol. 98 (1998)
11. Stefanowski, J.: The bagging and  $n^2$ -classifiers based on rules induced by MODLEM. In: *4th International Conference of Rough Sets and Current Trends in Computing*, vol. 3066, pp. 488–497 (2004)
12. Stefanowski, J.: On rough set based approaches to induction of decision rules. *Rough Sets Knowl. Disc.* **1**(1), 500–529 (1998)
13. Grzymala-Busse, J.W.: *LERS - a system for learning from examples based on rough sets*. *Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory*, vol. 11, pp. 3–18. Springer, The Netherlands (1992)
14. Stefanowski, J.: On combined classifiers, rule induction and rough sets. In: Peters, J.F., Skowron, A., Düntsch, I., Grzymala-Busse, J.W., Orłowska, E., Polkowski, L. (eds.) *Transactions on Rough Sets VI*. LNCS, vol. 4374, pp. 329–350. Springer, Heidelberg (2007)
15. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: *Proceedings of 8th IEEE International Conference on Computer Vision*, vol. 2, pp. 416–423 (2001)
16. Meilä, M.: Comparing clusterings—an information based distance. *J. Multivar. Anal.* **98**(5), 873–895 (2007)
17. Yang, A.Y., Wright, J., Ma, Y., Sastry, S.S.: Unsupervised segmentation of natural images via lossy data compression. *Comput. Vis. Image Underst.* **110**(2), 212–225 (2008)
18. Deng, Y., Manjunath, B.S.: Unsupervised segmentation of color-texture regions in images and video. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(8), 800–810 (2001)
19. An, N.Y., Pun, C.M.: Color image segmentation using adaptive color quantization and multiresolution texture characterization. *SIViP* **8**(5), 943–954 (2014)