Chilean Journal of Statistics
Vol. 2, No. 2, September 2011, 51–60

Spatial Statistics and Image Modeling
Special Issue: "III SEEMI"
Research Paper

Image segmentation based on wavelet feature descriptor and dimensionality reduction applied to remote sensing

Ricardo Dutra da Silva[1], William Robson Schwartz[1] and Helio Pedrini[1,*]

[1]Institute of Computing, University of Campinas, Campinas-SP, Brazil

(Received: 07 March 2011 · Accepted in final form: 21 June 2011)

Abstract

Image segmentation is a fundamental stage in several domains of knowledge, such as computer vision, medical applications, and remote sensing. Using feature descriptors based on color, pixel intensity, shape, or texture, it divides an image into regions of interest that can be further analyzed by higher level modules. This work proposes a two-stage image segmentation method that maintains an adequate discrimination of details while allowing a reduction in the computational cost. In the first stage, feature descriptors extracted using the wavelet transform are employed to describe and classify homogeneous regions in the image. Then, a classification scheme based on partial least squares is applied to those pixels not classified during the first stage. Experimental results evaluate the effectiveness of the proposed method and compares it with a segmentation approach that considers Euclidean distance instead of the partial least squares for the second stage.

Keywords: Image segmentation · Partial least squares · Wavelet transforms.

Mathematics Subject Classification: Primary 62H35 · Secondary 68U10.

1. Introduction

Image segmentation is a fundamental step in many image analysis tasks, including remote sensing, computer vision, and medical applications. The segmentation process consists of partitioning an image into a set of regions with similar features, in general described using texture, shape, gray level intensity or color information.

*Corresponding author. Helio Pedrini. Institute of Computing, University of Campinas, Campinas-SP, Brazil, 13083-852. Email: helio@ic.unicamp.br

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

Image segmentation approaches are commonly categorized into supervised and unsupervised. The former class of methods requires some prior knowledge, mainly regarding the number of classes (distinct patterns) present in an image, whereas the latter estimates the number of classes by automatically grouping image regions with a high degree of similarities. Although the lack of prior knowledge regarding the data classes makes the segmentation process complex, unsupervised algorithms are more flexible since they do not depend on the information estimated from a pre-determined segmentation set (training data set).

Several techniques have been proposed for image segmentation using region growing (see Deng and Manjunath, 2001), graph cuts (see Boykov and Funka-Leam, 2006; Rother et al., 2004), normalized cuts (see Shi and Malik, 2000), relaxation-based techniques (see Rosenfeld et al., 1976), neural network based approaches (see Shan et al., 2005), methods based on fuzzy theory (see Hall, 1992), level sets (see Vese and Chan, 2002), and Markov random fields (see Schwartz and Pedrini, 2007). Comprehensive reviews of image segmentation techniques are described by Pal and Pal (1993) and Cremers (2007).

This work presents a two-stage segmentation method. In the first stage, an initial segmentation and a clustering are computed to estimate parameters for modeling distinct classes in the image. Homogeneous regions of the image (those regions belonging to a single class) are labeled with the class identifier and not further considered. The class models are then used during the second stage to classify pixels labeled as heterogeneous in the first stage of the method. The modeling of the classes is obtained using a statistical tool known as partial least squares (PLS); see Wold (1985). The use of PLS has provided improvements on detection and recognition tasks; see Schwartz and Davis (2009), Schwartz et al. (2010) and Kembhavi et al. (2010). Its resulting projection vectors are able to discriminate among the multiple classes contained in the image. The use of a two-step approach enables the reduction of the computational cost significantly, since only a reduced number of pixels need to be considered during the second step (the most costly since feature descriptors are extracted for a pixel neighborhood). In addition, the first stage provides samples to build the PLS models, used to label pixels during the second stage.

This paper is organized as follows. Section 2 describes the wavelet transform, used to extract features, and the main concepts of PLS. In Section 3, the proposed method is presented and discussed. Experimental results obtained by applying the proposed method are shown in Section 4. Section 5 concludes with some final remarks.

2. Background

This section presents the main concepts of the wavelet transform and partial least squares (PLS). The wavelet transform is employed during the feature extraction to describe regions of the image. Then, after the estimation of an initial data clustering, PLS estimates multiple regression models in a one-against-all fashion.

2.1 Wavelet transform

The wavelet transform has similar properties to Fourier transform as a mathematical technique for signal analysis, the main difference between both is that wavelets are localized in both time and frequency, whereas the standard Fourier transform is only localized in frequency.

A signal can be decomposed by a wavelet transform through of a series of elementary functions, created from dilations and translations of a basis function ψ, which is known as the mother wavelet. The basis functions of a discrete wavelet transform, $\psi_{j,k}(t)$, of time independent variable t, can be expressed as

$$\psi_{j,k}(t) = 2^{-j/2}\, \psi(2^{-j}t - k),$$

where j and k guide the dilations and translations of the function ψ to generate a family of wavelets.

Wavelet transforms can be defined as a pair of lowpass and highpass filters represented by a sequence of coefficients; see Daubechies (1992), Mallat (1989) and Pun (2003). In a 2D wavelet decomposition, the filters are applied to an image in the horizontal and vertical directions, followed by a downsampling. The output of each level generates an image subband (LL) formed by the approximation coefficients and three subbands (LH, HL, HH) generated by the detail coefficients. The same process can be repeated on the LL image to generate the next decomposition level.

It has been shown that texture features can be extracted from wavelet coefficients in different frequency bands, since these coefficients show variations in horizontal, vertical and diagonal directions; see Unser (1995).

A well known feature based on wavelet coefficients is the energy, expressed as

$$E_{sb} = \sqrt{\frac{1}{m^2}\sum c(x,y)^2}, \tag{1}$$

where sb denotes the LL, LH, HL and HH subbands, $c(x,y)$ represents wavelet transform coefficients in the coordinates (x,y) for each one of these subbands containing $m \times m$ pixels. Wavelet energy reflects the distribution of energy along the frequency axis over scale and orientation, and it has been demonstrated to be very useful for texture characterization

2.2 Partial least squares

PLS is a method for modeling relations between sets of observed variables by means of latent variables. PLS estimates these variables as a linear combination of the original variables, represented by a matrix $\boldsymbol{X}$ (containing feature descriptors) and a vector $\boldsymbol{y}$ of response variables (usually, the class label). More details regarding the PLS method are described by Wold (1985), Elden (2004) and Rosipal and Kramer (2006).

PLS decomposes a zero-mean matrix $\boldsymbol{X}_{n\times m}$, composed by n samples and m feature descriptors, and a zero-mean vector $\boldsymbol{y}_{n\times 1}$ according to

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^\top + \boldsymbol{E} \quad \text{and} \quad \boldsymbol{y} = \boldsymbol{U}\boldsymbol{q}^\top + \boldsymbol{f},$$

where $\boldsymbol{T}$ and $\boldsymbol{U}$ are $n \times p$ matrices containing p extracted latent vectors, the $(m \times p)$ matrix $\boldsymbol{P}$ and $(1 \times p)$ vector $\boldsymbol{q}$ represent the loadings and the $n \times m$ matrix $\boldsymbol{E}$ and $n \times 1$ vector $\boldsymbol{f}$ are the residuals. Using the nonlinear iterative PLS algorithm (see Wold, 1985), a set of projection vectors is constructed and stored in the matrix $\boldsymbol{W} = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p)$, such that the following maximization is considered

$$[\mathrm{Cov}(\boldsymbol{t}_i, \boldsymbol{u}_i)]^2 = \max_{|\boldsymbol{w}_i|=1} [\mathrm{Cov}(\boldsymbol{X}\boldsymbol{w}_i, \boldsymbol{y})]^2,$$

where $\boldsymbol{t}_i$ is the ith column of matrix $\boldsymbol{T}$, $\boldsymbol{u}_i$ is the ith column of matrix $\boldsymbol{U}$ and $\mathrm{Cov}(\boldsymbol{t}_i, \boldsymbol{u}_i)$ is the sample covariance between latent vectors $\boldsymbol{t}_i$ and $\boldsymbol{u}_i$. After extracting the latent vectors $\boldsymbol{t}_i$ and $\boldsymbol{u}_i$, the matrix $\boldsymbol{X}$ and the vector $\boldsymbol{y}$ are deflated by subtracting their rank-one approximations based on $\boldsymbol{t}_i$ and $\boldsymbol{u}_i$. This process is repeated until the desired number of latent vectors has been extracted.

The regression coefficients $\boldsymbol{\beta}_{m\times 1}$ are estimated by

$$\boldsymbol{\beta} = \boldsymbol{W}(\boldsymbol{P}^\top \boldsymbol{W})^{-1}\boldsymbol{T}^\top \boldsymbol{y},$$

using the low dimensional representation obtained by the nonlinear iterative PLS algorithm. Thus, the regression response, y_v, for a feature vector $\boldsymbol{v}$ is obtained by

$$y_v = \overline{y} + \boldsymbol{\beta}^\top \boldsymbol{v},$$

where $\overline{y}$ is the sample mean of $\boldsymbol{y}$. The regression response is used during the segmentation to assign labels to pixels considered during the second stage, as described in Section 3.2.

3. Methodology

The developed image segmentation method is composed of two stages, as illustrated in Figure 1. The purpose of such division is to produce an initial segmentation of homogeneous regions and extract parameters for refinement of heterogeneous regions using PLS. In such a manner, computational cost is reduced while maintaining a proper segmentation of details.

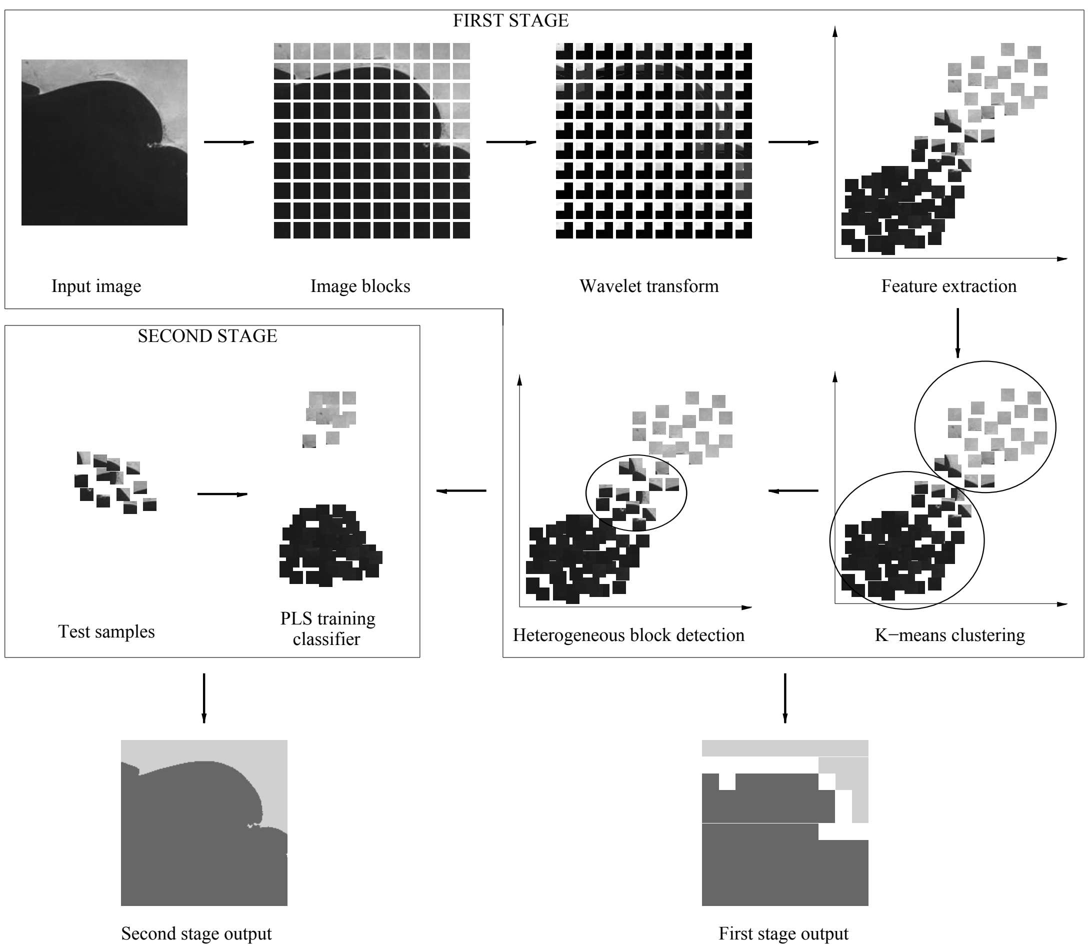


Figure 1. Proposed two-stage image segmentation method.

3.1 Initial segmentation

In the first stage, the input image is partitioned into square blocks with $m \times m$ pixels from which features are extracted to model distinct classes of an image and produce a first rough segmentation. The features are obtained from the subbands of a wavelet transform applied to each block.

In addition to energy, the use of some statistical measures is also common, such as the standard deviation. The smoothness expressed as

$$\mathrm{E}_{sb} = 1 - \frac{1}{1 + \sigma_{sb}^2},$$

assumes the value zero when the coefficients of a region are constant (smooth) and increases to one as the variance σ_{sb}^2 of subband sb becomes greater in rough regions For each image block, energy (see Equation (1)), standard deviation, and smoothness features are calculated for the subbands of the wavelet transform. In color images, the feature extraction process is applied to each band of the YCbCr color model. The k-means clustering algorithm is then used to group the feature vectors (blocks) into a set of classes.

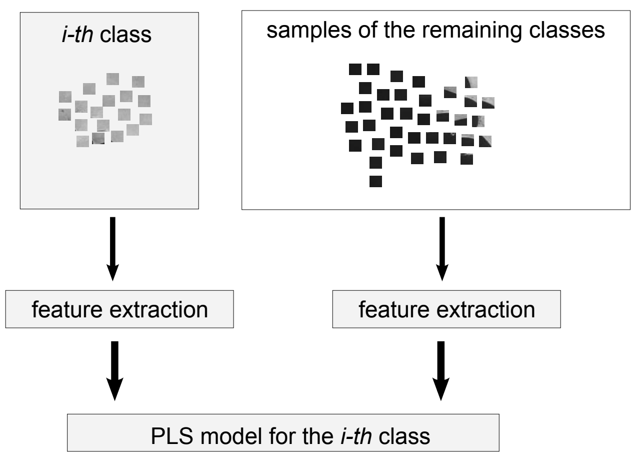


Figure 2. PLS learning considering a one-against-all classification scheme.

To reduce the blockiness effect, blocks located at boundaries are detected based on the similarity to the blocks in their class. The resulting measurement ranges from -1 to 1, indicating low to high similarity. The measure for a block b is defined as (see Kaufman and Rousseeuw, 1990)

$$\mathrm{Sim}(b) = \frac{\min\left(D(b,c)\right) - d(b)}{\max\left(\min\left(D(b,c)\right), d(b)\right)}, \tag{2}$$

where $D(b,c)$ is the mean distance from block b to the blocks of class c, and $d(b)$ is the mean distance from block b and the blocks of its own class. The inequality $|\mathrm{Sim}(b) - \mu_b| > t\,\sigma_b$ is then calculated, where t is a real value, μ_b and σ_b are respectively the mean and standard deviation of the similarities of b and the samples in its class. If the inequality is true, the block is marked as heterogeneous and its segmentation is refined in the final stage of the method, otherwise, all pixels of b are labeled and they are not further considered.

3.2 Final segmentation

The final segmentation is a pixelwise stage to classify the pixels remaining from the first stage. It is accomplished by using the partial least squares method. A one-against-all scheme, following the approach described by Schwartz et al. (2010), is used to learn PLS regression models. Each remaining pixel is assigned to that class which presents the highest regression response.

The procedure to learn models for multiple classes $C = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n\}$, where $\boldsymbol{c}_i$ represents exemplars of each class resulting from the k-means clustering, is illustrated in Figure 2. To build the PLS regression model for the ith class, the remaining samples $C \setminus \boldsymbol{c}_i$ are used as counter-examples of the ith class. The addition of remaining class samples as counter-example improves the discriminability of the PLS models; see Schwartz et al. (2009).

Once the PLS models have been built, the pixels remaining from the first stage are considered. For a given pixel, a feature vector is extracted from its neighborhood and projected onto each PLS model. Finally, the pixel is assigned to the class represented by the model with the highest regression response.

4. Results and Discussion

The proposed image segmentation method using PLS was compared to the previous work presented by da Silva et al. (2008), which employed Euclidean distance in the second stage of the method. The used block sizes were 4×4 and 8×8 pixels. The block size is an important parameter for the preservation of details and the identification of regions. The smaller the block, the better the detail identification. Nonetheless, the choice of larger blocks may result in better region description.

Wavelet transforms were applied using Symlets-2 filters with two levels of decomposition. Heterogeneous blocks that required further segmentation were identified with $t = 0.75$ in the first step of the segmentation (see Section 3.1) by using Equation (2). Value for t was experimentally obtained through the use of a number of validation images.

Table 1 presents some test images (in JPEG format) along with their dimensions and percentage of segmented pixels in the final stage. Since the refinement step is applied only to a smaller portion of the image, computational cost is significantly reduced.

Table 1. Test images and percentage of heterogeneous pixels in the final segmentation.

Images	Dimensions (pixels)	Segmented pixels in final stage (%)
Shark Bay	420×420	14.63
Moreno Glacier	340×340	20.75
Chesapeake Bay	512×512	13.67
Forest and Sand	512×512	5.85
Palm Island	512×512	14.81

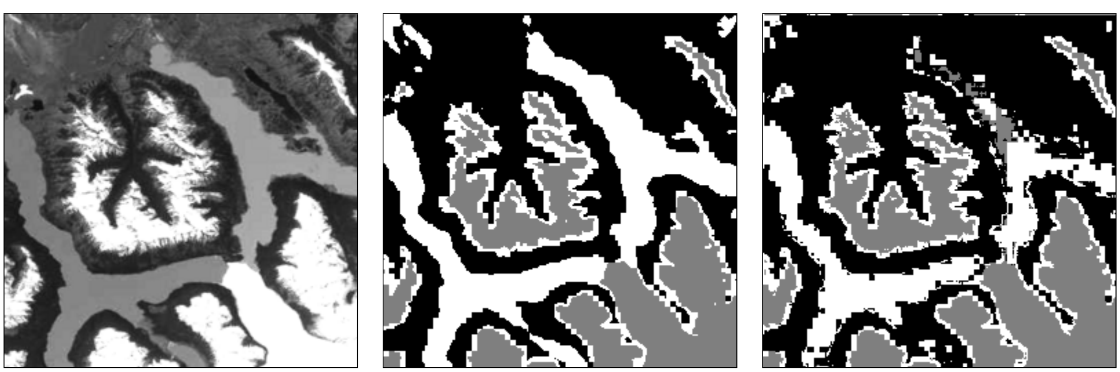
To evaluate the effectiveness of PLS compared to the previous approach (see da Silva et al., 2008), accuracy and kappa coefficient (see Liu et al., 2007) of the resulting segmentation were computed based on manually annotated ground truth images. Table 2 shows that the proposed method produced superior segmentation results for some images.

Table 2. Comparison between second stage with Euclidean distance and PLS.

Images	Euclidean		PLS	
	Accuracy	κ coefficient	Accuracy	κ coefficient
Shark Bay	0.970	0.930	0.972	0.934
Moreno Glacier	0.848	0.753	0.823	0.701
Chesapeake Bay	0.934	0.859	0.944	0.878
Forest and Sand	0.988	0.975	0.987	0.974
Palm Island	0.909	0.806	0.931	0.854

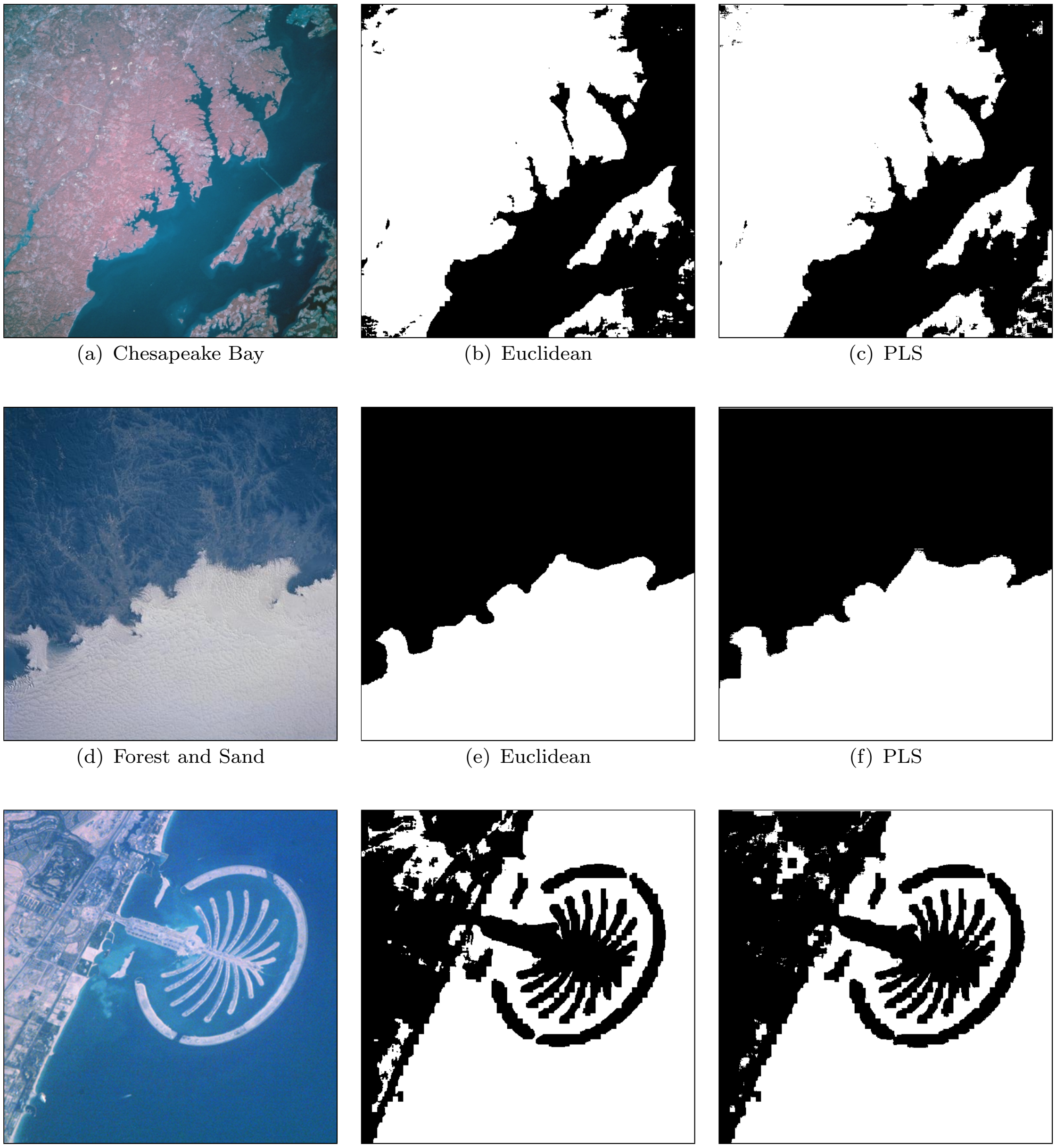
Figures 3 and 4 illustrate the results obtained with the methods based on Euclidean distance and partial least squares. Both methods preserved small and thin details, such as those present in Chesapeake Bay and Palm Island images. PLS-based image segmentation produced superior results, for instance, in land regions located at the lower right of the Chesapeake Bay image. However, it mis-segmented some parts of Moreno Glacier image. This fact is probably due to the reduced block size used in the experiments. Generally speaking, it can be observed that the method worked well in distinguishing regions of the images and it is suitable for remote sensing applications.

(a) Shark Bay (b) Euclidean (c) PLS



(d) Moreno Glacier (e) Euclidean (f) PLS

Figure 3. Segmentation results for two remote sensing images (to be continued in Figure 4).



(a) Chesapeake Bay (b) Euclidean (c) PLS

(d) Forest and Sand (e) Euclidean (f) PLS

(g) Palm Island (h) Euclidean (i) PLS

Figure 4. Segmentation results for three remote sensing images.

5. Conclusions

This paper described an image segmentation method of color textured remote sensing images based on wavelet transforms and partial least squares method. The use of partial least squares provided an improvement on the segmentation results when compared to a previous approach; see da Silva et al. (2008). As future investigation, different block sizes will be tested to demonstrate the effectiveness of the PLS-based image segmentation.

Acknowledgements

The authors would like to thank the Image Analysis Laboratory at NASA Johnson Space Center (see NASA, 2011), as well as CNPq and CAPES, from Brazil, for the financial support. This research was partially supported by FAPESP grant 2010/10618-3, Brazil.

References

Boykov, Y., Funka-Leam, G., 2006. Graph cuts and efficient N-D image segmentation. International Journal of Computer Vision, 70, 109-131.

Cremers, D., Rousson, M., Deriche, R., 2007. A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. International Journal of Computer Vision, 72, 195-215.

da Silva, R.D., Minetto, R., Schwartz, W.R., Pedrini, H., 2008. Satellite image segmentation using wavelet transforms based on color and texture features. International Symposium on Visual Computing, Las Vegas, USA.

Daubechies, I., 1992. Ten lectures on wavelets. CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM.

Deng, Y., Manjunath, B.S., 2001. Unsupervised segmentation of color-texture regions in images and video. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23, 800-810.

Elden, L., 2004. Partial least-squares vs. lanczos bidiagonalization-I: Analysis of a projection method for multiple regression. Computational Statistics and Data Analysis, 46, 11-31.

Hall, L.O., Bensaid, A.M., Clarke, L.P., Velthuizen, R.P., Silbiger, M.S., Bezdek, J.C., 1992. A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. IEEE Transactions on Neural Networks, 3, 672-682.

Kaufman, L., Rousseeuw, P.J., 1990. Finding Groups in Data An Introduction to Cluster Analysis. Wiley, New York.

Kembhavi, A., Harwood, D., Davis, L., 2010. Vehicle detection using partial least squares. IEEE Transactions on Pattern Analysis and Machine Intelligence, PP, 1.

Liu, C., Frazier, P., Kumar, L., 2007. Comparative assessment of the measures of thematic classification accuracy. Remote Sensing of Environment, 107, 606-616.

Mallat, S.G., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11, 674-693.

National Aeronautics and Space Administration, 2011. Earth from Space. `http://earth.jsc.nasa.gov/sseop/efs/`

Pal, N.R., Pal, S.K., 1993. A review on image segmentation techniques. PRIMUS, 26, 1277-1294.

Pun, C.-M., 2003. Rotation-invariant texture feature for image retrieval. Computer Vision and Image Understanding, 89, 24-43.

Rosenfeld, A., Hummel, R.A., Zucker, S.W., 1976. Scene labeling by relaxation operations. IEEE Transactions on Systems, Man and Cybernetics, 6, 420-433.

Rosipal, R., Kramer, N., 2006. Overview and recent advances in partial least squares. Lecture Notes in Computer Science, 3940, 34-51.

Rother, C., Kolmogorov, V., Blake, A., 2004. GrabCut: Interactive foreground extraction using iterated graph cuts. ACM Transactions on Graphics, 23, 309-314.

Schwartz, W.R., Davis, L.S., 2009. Learning discriminative appearance-based models using partial least squares. XXII Brazilian Symposium on Computer Graphics and Image Processing, pp. 322-329.

Schwartz, W.R., Guo, H., Davis, L.S., 2010. A robust and scalable approach to face identification. 11th European Conference on Computer Vision, Crete, Greece.

Schwartz, W.R., Kembhavi, A., Harwood, D., Davis, L.S., 2009. Human detection using partial least squares analysis. 12th International Conference on Computer Vision (ICCV), Kioto, Japan.

Schwartz, W.R., Pedrini, H., 2007. Color textured image segmentation based on spatial dependence using 3D co-occurrence matrices and Markov random fields. 15th International Conference in Central Europe on Computer Graphics, Visualization and Computer Visiom, Plzen, Czech Republic, pp. 81-87.

Shan, S., Sandham, W., Granat, M., Sterr, A., 2005. MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. IEEE Transactions on Information Technology in Biomedicine, 9, 459-467.

Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, 888-905.

Unser, M., 1995. Texture classification and segmentation using wavelet frames. IEEE Transactions on Image Processing, 4, 1459-1560.

Vese, L.A., Chan, T.F., 2002. A multiphase level set framework for image segmentation using the Mumford and Shah odel. International Journal of Computer Vision, 50, 271-293.

Wold, H., 1985. Partial least squares. In Kotz, S., Johnson, N.L., (eds.). Encyclopedia of Statistical Sciences. Wiley, New York, pp. 581-591.