REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
	Paper	

4. TITLE AND SUBTITLE

Segmenting Polarimetric SAR Images Using Robust Competitive Clustering

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dr. Paul Kersten; Dr. Roger Lee

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Air Warfare Center Aircraft Division
22347 Cedar Point Road, Unit #6
Patuxent River, Maryland 20670-1161

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Polarimetric Synthetic Aperture Radar (POLSAR) Images have great potential for land-use management, provided the images can be efficiently segmented. This paper describes the application of the robust competitive agglomeration (RCA) clustering algorithm to POLSAR images to segment the images. Examples are presented and future efforts are discussed.

15. SUBJECT TERMS

Polarimetric Synthetic Aperture Radar (POLSAR)

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Paul Kersten / Roger Lee
				5	19b. TELEPHONE NUMBER (include area code) (940) 397-4439 / (301) 342-0049

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

Segmenting Polarimetric SAR Images using Robust Competitive Clustering

Paul R. Kersten and Roger R-Y. Lee
Naval Air Warfare Center - Aircraft Division,
RF Sensors Branch, Patuxent River, MD 20670, USA
kerstenpr@navair.navy.mil, leery@navair.navy.mil

Abstract*
Polarimetric Synthetic Aperture Radar (POLSAR) Images have great potential for land-use management, provided the images can be efficiently segmented. This paper describes the application of the robust competitive agglomeration (RCA) clustering algorithm to POLSAR images to segment the images. Examples are presented and future efforts are discussed.

1. Introduction

J.S. Lee [1-2] has applied both hard c-means clustering (HCM) and fuzzy c-means clustering (FCM) to Synthetic Aperture Radar (SAR) images. Verdi et. al. [3] has also studied this approach on polarimetric high resolution tri-band SAR data and shown segmentation results for both the HCM and the FCM. The robust fuzzy c-means (RFCM) clustering algorithm has been applied to POLSAR images and also produced encouraging segmentation results [4]. This paper applies a modified version of the robust competitive agglomeration (RCA) clustering algorithm [5] to segment the POLSAR images. The RCA also provides an estimate for the number of clusters in the image. In section 2, POLSAR images are briefly described and previous publications in this area discussed. In section 3, a brief discussion of the applied version of the RCA is given. Section 4 gives some examples and section 5 contains the conclusions.

2. Polarimetric SAR Images

Polarimetric SAR images can be constructed from the complex scattering returns from the four possible polar combinations of transmit-receive returns of the radar: HH, HV, VH, and VV. Because of symmetry assumptions, HV and VH returns are identical yielding a 3-D complex scattering vector for each pixel in the image lattice. An incredible amount of preprocessing is required to form, register, and calibrate the image. The only feature used in this paper is the Coherence matrix, which is a Hermitian matrix defined as the outer product of three linear combinations of the complex scattering vectors: $HH+VV$, HV, and $HH-VV$. A real vector of dimension 9 is then constructed from the lower triangular part of this matrix. This feature vector, that is associated with each pixel of the image lattice, is used for clustering and classification. The dynamic range of this feature vector may be large and outliers are a frequent occurrence.

Du and Lee [2] applied FCM to segment SAR images using a distance measure based on Wishart measure, which replaced the usual Euclidean squared distance d_{ik}^2 in the FCM objective function. Similarly, RFCM replaces d_{ik}^2 with Huber's ρ function. Both approaches reduce the influence of outliers by replacing d_{ik}^2 with a slowly increasing function of distance.

3. Segmentation via Robust Clustering

Clustering is often used to segment images since segmentation is really pattern recognition, i.e. classifying each pixel [6-7]. After the pixel feature vectors are clustered into c distinct classes, labeling each pixel with the exemplar closest to it segments an image. The clustering method can employ either crisp sets as with the HCM or fuzzy sets as with the FCM, RFCM and RCA.

The HCM clustering algorithm is described in [8, p.55] and with the Wishart Measure in [1]. The FCM is a practical clustering algorithm that generalizes the HCM by replacing the class assignment with a membership vector whose elements represent the membership of the data point in each of the c distinct classes. The algorithm produces a fuzzy partition of the data and may be viewed as an unsupervised learning technique. The following description of the FCM is based on [8].

---

* Approved for Public Release; distribution is unlimited.

3.1 Fuzzy c-means

Consider N data samples forming the data set denoted by $X = \{x_1, x_2, \ldots, x_N\}$, where each sample $x_i \in R^P$. Assume that there are c classes and $u_{ik} = u_i(x_k) \in [0,1]$ is the membership of the *k-th* sample x_k in the *i-th* class v_i, where $v = (v_1, v_2, \ldots, v_c)$ is the set of exemplars or prototypes and $U = [u_{ik}]$ is the membership matrix. Each sample point x_k satisfies the constraint that $\sum_{i=1}^{c} u_{ik} = 1$. The FCM algorithm minimizes the function $J(U,v) = \sum_{k=1}^{N}\sum_{i=1}^{c} u_{ik}^{m_c} d_{ik}^2$ where $d_{ik} = \|v_i - x_k\|_2$ subject to the above constraint. The alternating optimization (AO) method is one technique to minimize $J(U,v)$. The power m_c of the membership is called the weighting exponent. A detailed version of this algorithm is given in [8, p.66]. HCM and FCM exemplars are linear statistics or weighted averages of the data points where the weights are scaled versions of the memberships. Unfortunately, linear statistics are known to be vulnerable to outliers [9]. HCM may also be viewed as a special case of FCM where the weighting exponent m_c is 1, and the data sample memberships in the classes are either 0 or 1. Compared to the FCM, the HCM is more often trapped in local minimum.

3.2. Robust Fuzzy c-Means clustering

To robustify the algorithm, a softer error function replaces the d_{ik}^2 term. One can replace d_{ik}^2 with $d_{ik} = \|v_i - x_k\|_1$ and then the resulting algorithm is called the fuzzy c-medians algorithm [10]. Another alternative is to replace d_{ik}^2 in the objective functional with Huber's ρ function. The objective function is $J(U,v) = \sum_{k=1}^{N}\sum_{i=1}^{c} u_{ik}^{m_c} \rho(d_{ik})$ where $d_{ik} = \|v_i - x_k\|_2$. The ρ function applied in the examples of section 4 is

$$\rho(x) = \begin{cases} \frac{1}{2}x^2, & if\ |x| \le 1 \\ |x| - \frac{1}{2}, & if\ |x| > 1 \end{cases}$$

whose form is quadratic when close to the exemplar, but linear when far from the exemplar. This particular ρ function is the one used by Huber in his early papers. The optimal memberships are then given by:

$$u_{ik} = 1 \Big/ \left[\sum_{j=1}^{c} \left(\frac{\rho(d_{ik})}{\rho(d_{jk})} \right)^{1/(m_c - 1)} \right]$$

The exemplars are computed by the weighted mean given by: $v_i = \sum_{k=1}^{N} u_{ik}^m w_{ik} x_{ik} \Big/ \sum_{k=1}^{N} u_{ik}^m w_{ik}$, where the Huber weights w_{ik} are dependent upon d_{ik} [11]. These estimates for v_i are W-estimators or robust recursive estimators because the weights w_{ik} are functions of v_i. The weights have the form $w(x) = \psi(x)/x$ where $\psi(x) = \rho'(x)$. In this case $w(x) = \begin{cases} 1, & |x| \le 1 \\ 1/|x|, & |x| > 1 \end{cases}$, which has the effect of gradually reducing the influence of the outliers. So, the exemplar v_i is a weighted combination of the sample values where the weights depend on both the membership of the k-th sample in the i-th cluster u_{ik} and a spatial Huber weight function [9].

The advantage of the RFCM clustering algorithm is its resistance to outliers, but at the expense of increased complexity in implementing the algorithm. For example, the W-estimator should be iterated at each stage of the RFCM, which of course would increase its time complexity by a factor proportional to the number of iterations. Moreover, a scaling constant is needed in the Huber weight, requiring an auxiliary estimate of dispersion. Here the auxiliary information is obtained from a robust estimator, the median absolute deviations about the median (MAD). Since the RFCM is non-linear in nature, it requires a better initialization for the exemplars. If an exemplar is too far from any data point, the membership of all the data points to this exemplar will be essentially zero, and the algorithm needs to deal with this special case to avoid underflow or overflow problems. In this paper, this problem is avoided by using Huber weights, which have infinite support yet vanishing weight. Finally, the most difficult part of the RFCM is that one must specify the number of clusters in advance, something that is usually not known. Determining the number of clusters is called the Validity problem. The RCA algorithm applied in this paper attempts to retain the clustering behavior of the RFCM and at the same time obtain a reasonable estimate of the optimum number of clusters.

The RCA minimizes the following functional

$$J(U,v)=\sum_{i=1}^{c}\sum_{k=1}^{N}u_{ik}^{2}\rho(d_{ik})-\alpha\sum_{i=1}^{c}\left[\sum_{k=1}^{N}u_{ik}w_{ik}\right]^{2}-\sum_{k=1}^{N}\lambda_{k}\left(\sum_{i=1}^{c}u_{ij}-1\right)$$

where the second term is a penalty function imposed to produce a parsimonious number of cluster. Note the weighting exponent m is now fixed at 2, which was assumed to simplify the analysis. The parameter α represents the tradeoff between the first two competing terms of the objective function $J(U,v)$. The first term is minimum when there are N clusters and the second term that is minimum when there is one cluster. Frigui [5] allows an initial period of agglomeration and then uses an exponential fader to reduce α to zero. The α and fading parameters must be set before running the program. In theory, one may grossly over-estimate the number of clusters and the RCA will converge to a "best" number of clusters. So now one does not need to specify the number of clusters in advance. However, one does have to specify three other parameters in advance, the α parameter, the fading parameter, and the cardinality threshold where one drops an active cluster. The RCA also uses the AO algorithm with the same centering statistics for the exemplars as the RFCM and a modified update to the memberships consisting of the sum of the RFCM membership plus a bias: $u_{ik}=u_{ik}^{rfcm}+\beta_{ik}$. The bias term is given by

$$\beta_{ik}=\frac{\alpha}{\rho(d_{ik})}\left(N'_{ik}-\overline{N}_{ik}\right) \text{ where}$$

$$N'_{ik}=w_{ik}\sum_{j=1}^{N}u_{ij}w_{ij}=w_{ik}N_{i} \quad \text{and} \quad \overline{N}_{ik}=\frac{\sum_{i=1}^{c}\frac{N'_{ik}}{\rho(d_{ik})}}{\sum_{i=1}^{c}\frac{1}{\rho(d_{ik})}}.$$

The bias term may be thought of as the correction to the RFCM for penalizing the increase in the number of clusters.

4. Examples

The set of parameters for the RCA is large enough to make it very adaptable to different image environment, but this also means that one has to initialize these parameters. A means of automatically setting them will be needed if this algorithm is ever fielded; however, for this paper one can only discuss the influence of each of these parameters and present some examples of these images. The parameter α determines the trade-off between the within cluster error and the excess cluster number penalty. The setting of α and the fading parameter provide the trade-off of fidelity and computational complexity. Fewer clusters reduce the time and space complexity of the algorithm at the expense of losing the finer granularity in objects in the image. At each iteration, the ratio of the first term to the second is computed to maintain the same relation between these two competing terms. Then $\alpha=\eta\cdot ratio\cdot\exp(-index/\tau)$ where *index* is the iteration number, η is the constant that controls the relative importance of the two terms, $\exp(-index/\tau)$ is the exponential fader and τ is a time constant. After $index/\tau>3$ the effect of the second term has essentially faded out.

The granularity of the objects that can be resolved is also a function of another parameter, which discards an active cluster once it falls below a fuzzy cardinality threshold. Multiple cluster exemplars sharing a given cluster of points have still been observed where both clusters were above cardinality threshold and therefore were stable. Also influencing the number of clusters is the scale parameter used in the Huber ρ function. If not chosen too small, a given cluster will be broken-up into a cluster of clusters, which can be either interpreted as a modeling plus or minus. Discovering the correct set of parameters for a given set of POLSAR images is time-consuming and also depends upon what information one tries to extract from the image. These parameters maybe differ from one collection of images to another. Considerable effort remains to be invested in learning these tradeoffs.

The images presented in this section apply to land-use management, where one is trying to classify large physical features and thereby emphasizing parsimony in cluster number over within cluster error. Figure 1 is an image of blueberry fields, where one is interested in classifying harvested and non-bearing blueberry fields versus bearing blueberry fields.

Figure 1 is a 341x341 pixel POLSAR span image where the pixel reflects the total power in the HH, HV, and VV returns after properly censoring and scaling. The image consist of various patches of blueberry fields in different stages of development. The lighter fields are ready to harvest and the darker fields are either non-bearing or harvested. Figure 2 shows the image constructed from 6 clusters derived from the RCA algorithm, with an initial set of 20 clusters, after 50 iterations. Here τ is 10, η is 4, the scale factor is three times the MAD estimator and the minimum cardinality was set at one 50^{th} of the number of pixels in the image or 1162 pixels. The RCA reduces the number of clusters by a factor of at least three, which after 50 iterations represents five time

constants, i.e. $index = 5\tau$ so the influence of the second term has been essentially discounted.

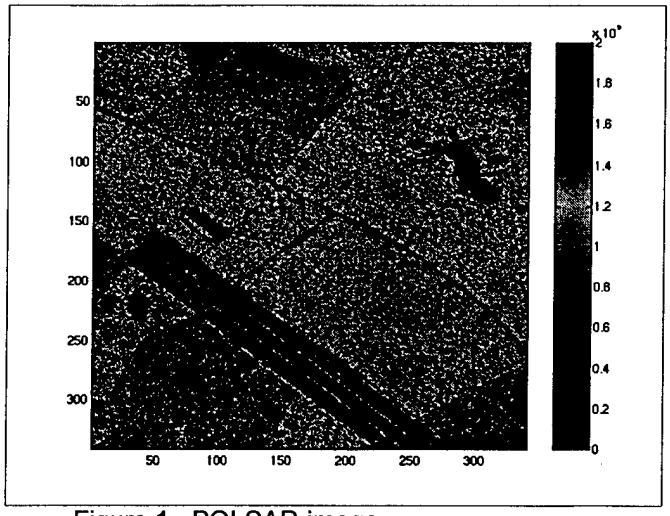

Figure 1. POLSAR image.

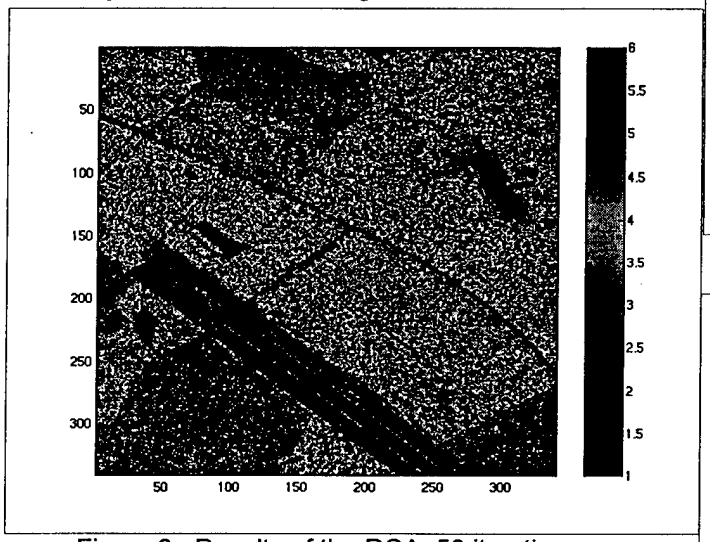

Figure 2. Results of the RCA, 50 iterations.

The six exemplars generated from the RCA give some insight into representation of the image features as exemplars and into the capability of the RCA to aid in exploratory data analysis. Figure 3 shows the plots of the six exemplar vectors, although one immediately notices that there only appears to be five. The exemplars are ordered in increasing energy and two most energetic exemplars are very similar so they appear to be the same on this scale. What is nice about this concise representation of this image is that one essentially assign exemplars or groups of exemplars to specific features in the image. The lowest energy exemplar represents the ground, the second exemplar represents the harvested blueberries. The third and fourth largest exemplars represent the blueberries ready for harvest and finally the fifth and sixth exemplars represent groups of trees or larger structure elements in the blueberry fields. What is even more revealing in this plot is the possibility of data reduction. Since only the 1st, 2nd, 3rd, 6th and 9th elements of the feature vector seem to contain discriminatory information, one can drop the dimensionality of the representation by nearly one-half with its accompanying reduction in time- and space-complexity..

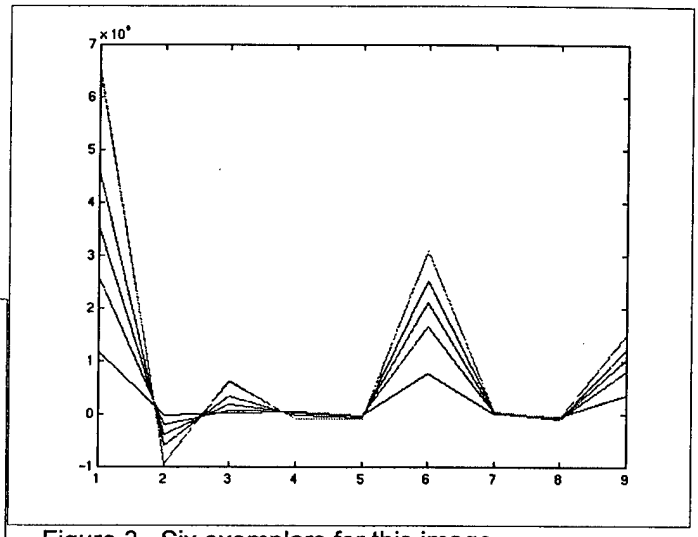

Figure 3. Six exemplars for this image.

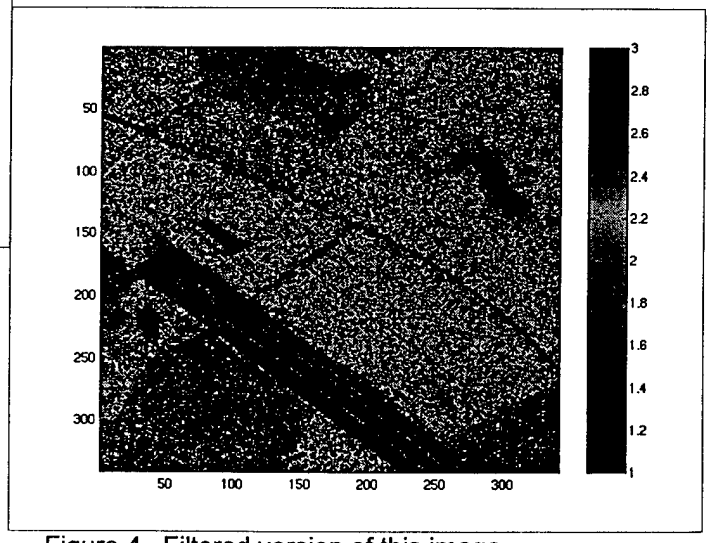

Figure 4. Filtered version of this image.

Figure 4 contains a filtered version of figure 2 where the first two low energy clusters have been mapped to a first value, third and fourth clusters have been mapped to a second value and the two most energetic clusters have been mapped to a third value. The resulting image tends to segments the yielding

blueberry field. Although the segmentation is not totally distinct, it must be remembered that nature of the growths does not produce the distinct boundaries characteristic of manmade objects, so the segmentation is not as distinctive as objects produced on an assembly line.

5. Conclusions

The modified robust competitive agglomeration (RCA) clustering algorithm has been applied to segment POLSAR images. Ordering the clusters by received signal energy, one can further segment the image structures with the low energy clusters representing low growth vegetation and the high energy clusters representing large growth structures and man-made objects. Outliers and exceedingly small structures are removed from the images by setting the fuzzy cardinality for rejection. The RCA also is a good algorithm for exploratory data analysis tool because one can drive the clustering to emphasize either fidelity or simplicity. Simplicity produces a parsimonious representation and a more efficient algorithm. Fidelity allows the user to adjust the granularity to match the application. In either case, the RCA suggests which data vector components are contributing to the discrimination and thus which components may be eliminated to reduce the dimensionality without loss of discrimination.

6. References

[1] J. S. Lee, M. R. Grunes, T. L. Ainsworth, L. Du, D. L. Schuler and S. R. Cloude, "Unsupervised Classification Using Polarimetric Decomposition and Complex Wishart Classifier," Proceedings IGARSS 98, Seattle, WA, pp. 2178-2180, 1998 July 6-10.

[2] L. Du and J.S. Lee, "Fuzzy classification of earth terrain covers using complex polarimetric SAR data, Int. J. Remote Sensing, 1996, Vol. 17, No 4, pp. 809-926.

[3] J. Verdi, R, Lee, P. R. Kersten, et. al, "Comparative Analysis of Polarimetric SAR Image Exploitation Algorithms for Environmental Stress-Change Monitoring," IGARSS Proceedings, Vol. V., 28 June - 2 July 1999, pp 2449-2451.

[4] P. Kersten, R. Lee, J. Verdi, R. Carvalho, S. Yankovich, "Segmenting SAR Images with Fuzzy Clustering," NAFIPS2000 Conference, Atlanta, GA, July 13-15, 2000 , p.105-108.

[5] H. Frigui, and R. Krishnapuram, "A Robust Competitive Clustering Algorithm With Applications in Computer Vision" , IEEE Transactions on Pattern Analyis & Machine Intelligence, Vol. 10, No. 5, May 1999

[6] J. Bezdek and S. Pal, Fuzzy Models for Pattern Recognition, Bezdek and Pal Editors, IEEE Press, NY, 1992

[7] J.C. Bezdek and M. M. Trivedi, Low level segmentation of aerial images with fuzzy clustering, IEEE Trans. Syst. Man. Cyb. SMC-16, 589-598 (1986).S. R. Cloude, E. Pottier, "An Entropy Based Classification Scheme for Land Applications of Polarimetric SAR," IEEE Trans. GRS, Vol. 35, No. 1, pp. 68-78, January 1997.

[8] J. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981

[9] D.C. Hoagen, F. Mosteller, and J.W. Tukey, *Understanding Robust and Exploratory Data Analysis*, Wiley, New York, 1983, p. 367.

[10] Kersten, P., "Fuzzy Order Statistics and Its Application to Fuzzy Clustering, IEEE Transactions on Fuzzy Systems,Vol 7, No 6, Dec 1999, pp 708-712.

[11] Y Choi and R. "Krishnapuram, "Fuzzy and Robust Formulations of Maximum-Likelihoor-Based Gaussian Mixture Decomposition,", Fifth IEEE International Conference on Fuzzy Systems, Sept 8-11, New Orleans, LA, Vol 3, p. 1899 -1905.