Synthetic Aperture Radar Feature Selection for Dual Polarized ScanSAR Data

N. Gökhan Kasapoğlu
Dept. of Electronics and Communication Engineering
Istanbul Technical University
Istanbul, Turkey
e-mail: gokhank@itu.edu.tr

Abstract— **Synthetic aperture radar (SAR) ScanSAR data has advantages on oceanographic remote sensing applications regarding its large coverage and sufficient resolution. However terrestrial downlink bandwidth is limited and therefore up to dual polarized (e.g., HH and HV) ScanSAR data can be achieved today's spaceborn systems (e.g., RadarSAT-2). In this study grey level co-occurrence matrix was employed to extract SAR features for both HH and HV channels. Additionally some of band math products such as HH/HV and HH-HV were used as candidate SAR features. Selection of optimum SAR features is crucial and application dependent. In this study, selection strategies based on SAR data assimilation was introduced and relation of conventional separability criterions on SAR data assimilation and pattern recognition were discussed.**

Keywords- *SAR Feature Selection, Discrimination Analysis, Data Analysis Fusion, Data Assimilation.*

I. INTRODUCTION

Synthetic aperture radar (SAR) data is an effective tool in remote sensing applications especially for sea ice analysis. Electromagnetic waves penetrate clouds therefore cloud cover does not affect monitoring. Oceanographic applications such as sea ice analysis and wind retrieval requires large coverage area with reasonable resolution. RadarSAT-2 ScanSAR data cover these requirements therefore it is mostly used by ice services. However its incidence angle range changes between 19 deg and 50 deg. which cause some ambiguities in sea ice analysis.

ScanSAR data can be achieved up to dual polarized due to the limited terrestrial downlink band width. Standard SAR products contain 8 bits digital numbers (DNs) which are from 0 to 255. Incidence angle correction has already applied for standard SAR products. Using sigma nought is physically more meaningful than DNs. HH and HV sigma nought can be used for sea ice analysis as tonal SAR features and, texture feature extraction methods [1-3] can be applied to extract textural SAR features for every individual tonal features.

Selection of optimum SAR features is crucial to reach improved sea ice analysis for operational applications. In general selection can be applied based on a selection measure which depends on application. Separability measures widely used as selection criterion in pattern recognition applications such as classification [4]. For data assimilation, a performance criterion for data assimilation can also be used to select appropriate SAR features [5]. In general, analysis bias and standard deviation are employed to measure performance of the data assimilation method [6]. Therefore in this study analysis bias based selection criterion were utilized in the proposed selection method. Normally an exhaustive test of all possible SAR feature combinations can be done to verify the selection algorithm. But testing all possible SAR feature combinations is very time consuming, therefore top-down and bottom-up strategies were used to specify SAR feature combinations. Additionally existence of sufficiently correlated SAR features can always prevent selecting optimal SAR feature set. Therefore selected features should be independent from each other.

Incidence angle dependency is very crucial in SAR based sea ice analysis. There may not be an optimum SAR feature available through the whole range direction for every individual incidence angle. Therefore selection of SAR features for every individual incidence angle is essential to reach robust results.

In this study reasonable number of SAR feature were selected both whole range direction and every individual incidence angle. In section II SAR feature extraction and candidate SAR features were introduced, in section III conventional separability measures and discrimination analysis were discussed. Performance criterion of the data assimilation method was employed as separability measure in section IV and conclusion were made in section V.

II. CANDIDATE SAR FEATURES

Single polarized HH ScanSAR data in C band (5.3 GHz) has been used successfully in operational sea ice analysis for many years. However wind open water interaction causes some ambiguities in HH SAR data. Availability of dual polarized (i.e., HH and HV) Radarsat-2 ScanSAR data diminish these ambiguities. Better discrimination between open water and sea ice can be obtained by using HV data. However signal level of HV data is lower than HH and therefore new ice detection is more problematic in HV data use than HH. Nevertheless incidence angle dependency is weaker in HV than HH. Radar backscatter (i.e., sigma nought) which is relevant to radar cross section is more meaningful to use as tonal SAR features for both HH and HV data than using

978-1-4244-9616-7/11/$26.00 ©2011 IEEE

TABLE I
EXTRACTED SCANSAR FEATURES

F. ID	UNIT	POLARIZATION	DESCRIPTION
0	AMPLITUDE SIGMA NOUGHT	HH	SIGMA NOUGHT
1			LEE FILTERED IMAGE
2			MEAN
3			VARIANCE
4			HOMOGENEITY
5			CONTRAST
6			DISSIMILARITY
7			ENTROPY
8			SECOND MOMENT
9			CORRELATION
10			DATA-RANGE
11			MEAN EUCLIDEAN DISTANCE
12		HV	SIGMA NOUGHT
13			LEE FILTERED IMAGE
14			MEAN
15			VARIANCE
16			HOMOGENEITY
17			CONTRAST
18			DISSIMILARITY
19			ENTROPY
20			SECOND MOMENT
21			CORRELATION
22			DATA-RANGE
23			MEAN EUCLIDEAN DISTANCE
24		COMBINATION OF POLARIZATIONS	HH/HV
25			HH-HV
26	SIGMA NOUGHT POWER	HH	SIGMA NOUGHT
27		HV	SIGMA NOUGHT
28	SIGMA NOUGHT POWER (dB)	HH	SIGMA NOUGHT (dB)
29		HV	SIGMA NOUGHT (dB)

standard SAR products (DNs). Human visual system (manual sea ice analysis) is very successful in sea ice analysis. The reason might be textural detection capability. In this study texture based SAR features were extracted from HH and HV sigma nought to reach additional and valuable information of sea ice.

Grey level co-occurrence based feature extraction method were utilized to extract texture features from HH and HV ScanSAR data. 64 grey levels with sub image size of 9x9 pixels were chosen to extract SAR features [3]. Additionally some band products such as HH/HV, HH-HV and HH and HV sigma nought in different units such as dB and power were added in to the candidate SAR feature set [6]. The candidate ScanSAR feature set was depicted in Table I.

III. SEPARABILITY MEASURES AND DISCRIMINATION ANALYSIS

Class separability is very crucial for selection of SAR features in pattern recognition. Several separability measures exist in the literature. Every separability measure has its own statistical assumption and this causes various results. Additionally existence of sufficiently correlated features can always prevent selecting the optimal feature set.

Between and within class scattering matrix based measures are widely used for separability analysis. The within class scatter matrix is related to compaction of the class. The within class scatter matrix is defined as

$$S_w(k) = \sum_{i=1}^{k} \sum_{x \in C_i} (x-\mu_i)(x-\mu_i)^T \tag{1}$$

where k is the class number, μ_i is the mean vector of the C_i class. The between class scatter matrix measures the separation between classes and is defined as follows:

$$S_B(k) = \sum_{i=1}^{k} n_i(\mu_i - \mu)(\mu_i - \mu)^T \tag{2}$$

where n_i is the number of samples in class C_i and μ is the mean of all samples. Two different measures were described based on within class and between classes scattering matrix as follows:

$$d_1 = tr(S_w S_b)^{-1} \tag{3}$$

$$d_2 = \frac{tr(S_b)}{tr(S_w)} \tag{4}$$

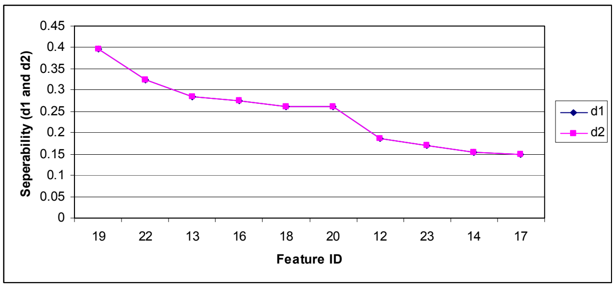

Fig. 1. The best ten discriminative SAR features based on separability measures d_1 and d_2.

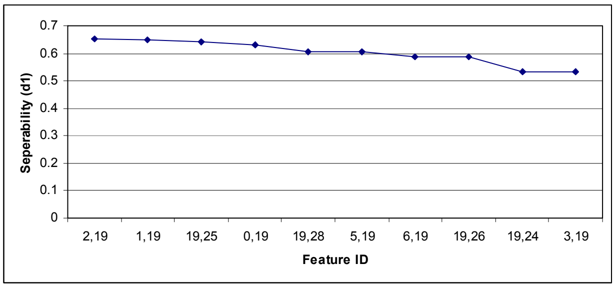

Fig. 2. The best ten discriminative combinations of two SAR features based on separability measure d_1.

The best ten discriminative SAR features and combination of two SAR features were depicted in Fig. 1 and Fig. 2 for separability measures d1 and d2 respectively. It is obvious that some of SAR features shown in the candidate SAR features table (see Table 1) are strongly dependent each other. For example such dependent features are: Sigma nought (F.ID: 0), lee Filtered image (F.ID: 1), Mean (F.ID: 2) and Contrast (F.ID: 5). Separability values of F.IDs 2-19, 1-19 and 0-19 are very close to each other (see Fig. 2).

Divergence analysis and transform divergence are widely used as separability measures in the literature. Divergence of pair of normal distributions [4] was given as follows:

$$d_3 = d_{ij} = \frac{1}{2} Tr\left\{\left(\Sigma_i - \Sigma_j\right)\left(\Sigma_j^{-1} - \Sigma_i^{-1}\right)\right\} + \frac{1}{2} Tr\left\{\left(\Sigma_i^{-1} + \Sigma_j^{-1}\right)\left(m_i - m_j\right)\left(m_i - m_j\right)^T\right\} \quad (5)$$

where Σ_i and m_i are covariance matrix and mean of class i. In this study separability analysis results were obtained for sea ice and open water. Additional separability measure based on divergence such as Transform divergence was also used in the experiments. Transform divergence were given as follows:

$$d_4 = d_{ij}^T = 2(1 - e^{-d_{ij}/8}) \quad (6)$$

The best ten discriminative SAR features and the combination of two SAR features were shown in Fig. 3 and

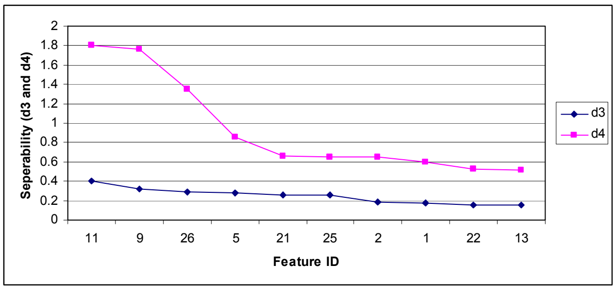

Fig. 3. The best ten discriminative SAR features based on separability measures divergence (d_3) and the transform divergence (d_4).

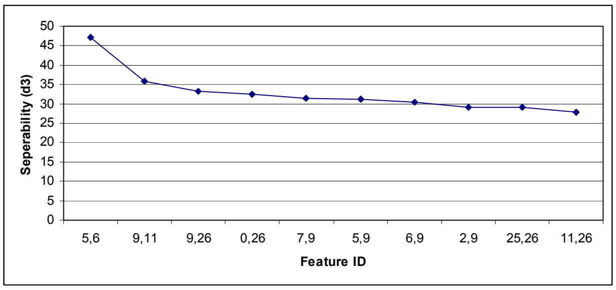

Fig. 4. The best ten discriminative combinations of two SAR features based on the divergence (d_3).

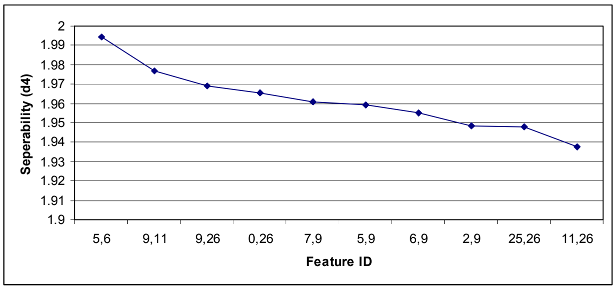

Fig. 5. The best ten discriminative combinations of two SAR features based on the transform divergence (d_4).

Figures 4, 5 based on the divergence (*d3*) and the transform divergence (*d4*) analysis respectively.

The Bhattacharyya distance (BD) is another discrimination measure and has been employed as projection index in projection pursuit (PP) algorithm [7] which can be used to select valuable bands in hyperspectral data. Bhattacharyya distance was given as follows:

$$BD = \frac{1}{8}\left(m_i - m_j\right)^T \left[\frac{\Sigma_i + \Sigma_j}{2}\right]^{-1} \left(m_i - m_j\right) + \frac{1}{2} \ln \left[\frac{\left|\frac{\Sigma_i + \Sigma_j}{2}\right|}{\sqrt{\left|\Sigma_i\right|\left|\Sigma_j\right|}}\right] \quad (7)$$

Jeffries-Matusita (JM) distance is a feature similarity measure and can be integrated during the SAR feature selection process. JM distance can be defined as follows:

$$JM = J_{ij} = 2\left(1 - e^{-BD}\right) \quad (8)$$

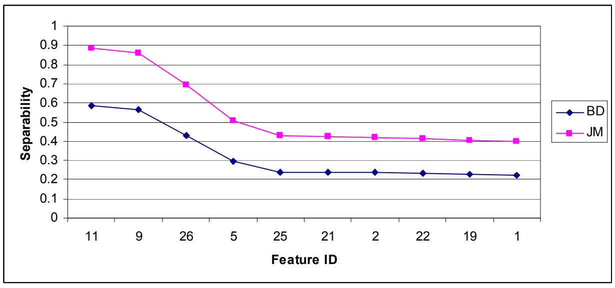

Fig. 6. The best ten discriminative SAR features based on separability measures Bhattacharyya distance (BD) and Jeffries Matusita (JM).

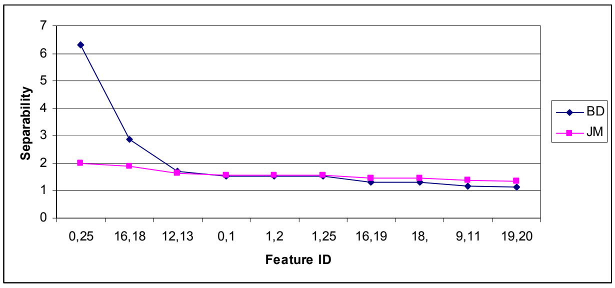

Fig. 7. The best ten discriminative combinations of two SAR features based on the Bhattacharyya distance (BD) and Jeffries Matusita (JM).

Discrimination analysis results based on BD and JM distance were given in Fig. 6 and Fig. 7 for the best ten discriminative SAR features and the combination of two SAR features respectively.

IV. SAR Feature Selection Based on Analysis Bias

Data analysis fusion is a novel method to combine analysis results from various information sources [8]. Passive microwave and SAR data analysis fusion were introduced in [6]. Analysis bias is employed as a performance criterion in SAR data analysis fusion. Manual analysis charts which produced by image analysis experts were used for validation [9]. Small analysis bias is desired. Analysis bias was utilized as a SAR feature selection criterion. All possible SAR feature combinations should be checked to find optimum SAR features. However this kind of analysis is very time consuming. Therefore in this study top-down and bottom-up strategies were applied to find valuable SAR features.

For top-down strategy; leave every candidate SAR feature out and make analysis with the rest of the SAR features and calculate analysis bias. When the minimum analysis bias is obtained, redundant SAR feature which is out is specifies. Same SAR feature reduction procedure is applied until to reach desired number of SAR features. Obtained combination of best four, three and two SAR features for HH and HV data were given in Table II and Tale III.

For bottom-up strategy; every individual candidate SAR features is employed to make analysis one by one. When the minimum analysis bias is obtained, the best SAR feature is specifies. Then rest of the SAR features are added to the best feature set one by one and analysis is performed with

Table II
The Best Combinations of HH SAR Features on Top Down Strategy

Number Of Features	Feature ID	Out Feat. ID	Analysis Bias	Background Bias
4	3 – 4 – 6 – 10	2	-0.02560	**-0.02664**
	2 – 4 – 6 – 10	3	-0.02521	
	2 – 3 – 6 – 10	4	**-0.02497**	
	2 – 3 – 4 – 10	6	-0.02543	
	2 – 3 – 4 – 6	10	-0.02517	
3	3 – 6 – 10	2	-0.02560	
	2 – 6 – 10	3	-0.02519	
	2 – 3 – 10	6	-0.02542	
	2 – 3 – 6	10	**-0.02515**	
2	3 – 6	2	-0.02580	
	2 – 6	3	**-0.02539**	
	2 – 3	6	-0.02561	
1	6	2	-0.02607	
	2	6	**-0.02587**	

Table III
The Best Combinations of HV SAR Features Based on Top Down Strategy

Number Of Feat.	Feature ID	Out Feat. ID	Analysis Bias
4	20 – 21 – 22 – 23	15	-0.02733
	15 – 21 – 22 – 23	20	-0.02658
	15 – 20 – 22 – 23	21	**-0.02704**
	15 – 20 – 21 – 23	22	-0.02729
	15 – 20 – 21 – 22	23	-0.02726
3	21 – 22 – 23	15	-0.02672
	15 – 22 – 23	21	**-0.026427**
	15 – 21 – 23	22	-0.026651
	15 – 21 – 22	23	-0.026637
2	22 – 23	15	-0.026567
	15 – 23	22	-0.026480
	15 – 22	23	**-0.026471**
1	22	15	-0.026611
	15	22	**-0.026505**

the best (selected) and added SAR feature. Similar way most valuable SAR feature is specified and added to the best feature set. The procedure is applied until to reach desired number of SAR feature. Obtained combination of best four, three and two SAR features from relatively 12 independent SAR features was depicted in Table IV.

V. Conclusions

Optimum SAR feature selection is very crucial to reach improved results in SAR based automated sea ice analysis [5,6]. Selection procedure should utilize application dependent selection criterion. Conventional discrimination analysis is also important but results can be used mostly in conventional pattern recognition applications such as classification. However Jeffries Matusita distance which is

TABLE IV
THE BEST COMBINATIONS OF SAR FEATURES BASED ON BOTTOM UP STRATEGY

NUMBER OF FEATURES	FEATURE ID	SELECTED FEATURE IDS	ANALYSIS BIAS
1	**2**	√	**-0.025879**
	4		-0.026626
	7		-0.026902
	8		-0.027099
	9		-0.026848
	11		-0.027967
	14		-0.027378
	16		-0.027816
	19		-0.027552
	20		-0.027335
	21		-0.026825
	23		-0.026618
2	2 – 4		-0.025883
	2 – 7		-0.026147
	2 – 8		-0.026338
	2 – 9		-0.026117
	2 – 11		-0.027201
	2 – 14		-0.026626
	2 – 16		-0.027078
	2 – 19		-0.026813
	2 – 20		-0.026590
	2 – 21		-0.026064
	2 – 23	√	**-0.025875**
3	**2 – 23 – 4**	√	**-0.025876**
	2 – 23 – 7		-0.026135
	2 – 23 – 8		-0.026323
	2 – 23 – 9		-0.026114
	2 – 23 – 11		-0.027175
	2 – 23 – 14		-0.026595
	2 – 23 – 16		-0.027038
	2 – 23 – 19		-0.026771
	2 – 23 – 20		-0.026560
	2 – 23 – 21		-0.026052
4	2 – 23 – 4 – 7		-0.026137
	2 – 23 – 4 – 8		-0.026326
	2 – 23 – 4 – 9		-0.026119
	2 – 23 – 4 – 11		-0.027172
	2 – 23 – 4 – 14		-0.026586
	2 – 23 – 4 – 16		-0.027029
	2 – 23 – 4 – 19		-0.026756
	2 – 23 – 4 – 20		-0.026552
	2 – 23 – 4 – 21	√	**-0.026052**

feature similarity measure can be integrated during the selection of most independent features. Then analysis bias based selection criterion can be used with bottom up strategy to find optimum SAR features for SAR data assimilation for sea ice analysis.

REFERENCES

[1] R.M. Haralick, K. Shanmugam and I. Dinstein, "Textural features for image classification," IEEE Trans. Systems, Man and Cybernetics, vol. 3, no.6, pp. 610-621, Nov. 1973.

[2] N. G. Kasapoglu, I. Papila, S.Kent and B. Yazgan, "Segmentation and Classification of SAR Data with Co- Occurrence Matrix for Texture Features", s. 717, 4th European Conference on Synthetic Aperture Radar (EUSAR'2002), Cologne, Germany.

[3] T. Bovith and S. Andersen, "Sea ice concentration from single-polarized SAR data using second-order grey level statistics and learning vector quantization," Scientific Report 05-04, Danish Meteorological Institute, Copenhagen 2005.

[4] J. A. Richards and X. Jia, Remote Sensing Digital Image Analysis, 4th ed., pp. 267-294, *Springer-Verlag* Berlin Heidelberg, 2006.

[5] N. G. Kasapoglu, T. Carrieres and M. Buehner, "Modelling of synthetic aperture radar features for sea-ice analysis and forecasting," the Third Radarsat-2 Workshop, Saint-Hubert, Canada, Sept. 2010.

[6] N. G. Kasapoglu, "Synthetic aperture radar and passive microwave analysis fusion", Recent Advances in Space Technologies (RAST), Istanbul, Turkey, June 2011.

[7] L.O. Jimenez, A.M. Morell and A. Creus, "Classification of hyperdimensional data based featre and decision fusion approaches using projection pursuit, majority voting, and neural networks," *IEEE Trans. Geosci. Remote Sensing,* vol. 37, no. 3, May 1999.

[8] A. Caya, M. Buehner and T. Carrieres, "Analysis and forecasting of sea ice conditions with three dimensional variational data assimilation and a coupled ice-ocean model," *Journal of Atmospheric and Oceanic Technology*, pp. 353-369, Feb. 2010.

[9] T. Carrieres, B. Greenan, S. Prinsenberg and I.K. Peterson, "Comparison of Canadian ice charts with surface observations off Newfoundland, winter 1992," *Atmosphere-Ocean*, vol. 34, pp. 207–236, 1996.