A Multi-Variate Contour Detector for High-Resolution Polarimetric SAR Images

D. Borghys, C. Perneel and M. Acheroy
Royal Military Academy
Signal & Image Centre
Av. de la Renaissance 30, B-1000 Brussels
Dirk.Borghys@elec.rma.ac.be

Abstract

This article presents a new method for contour detection in high-resolution polarimetric SAR images. The method is based on multi-variate statistics of speckle in homogeneous regions in a SAR image and uses a hypothesis test for the difference in variance between two adjacent regions to find the contours. The detector is directly applied on the single-look complex polarimetric SAR image. A pre-processor, which is also based on multi-variate statistics, is used to focus the attention of the detector on potentially non-homogeneous regions within the image. Results of applying the contour detector on L-band polarimetric SAR images are also presented.

1. Introduction

The presented work is part of a project that aims at automatically register SAR images with other types of images for remote sensing applications. The general idea is to use a map as an aid for the registration. On a map of the same region as the images the main objects (roads, forests, villages, etc.) are detected automatically. Then, for each type of image, some of the objects found on the map are also detected. Matching the objects will provide a first registration. The purpose of the contour detector presented in this article is thus to automatically find edges between large uniform regions that are likely to be present on maps.

Automatic contour detection in Synthetic Aperture Radar (SAR) images is a difficult problem due to the presence of speckle. Classical methods [3, 5] give unsatisfactory results. However, the statistics of speckle in homogeneous regions of SAR images can be accurately modeled and some methods that exploit these statistics to detect contours have been presented in literature [2, 4, 8, 9]. These methods were mainly applied on multi-look low-resolution images.

Thanks to the increased spatial resolution of the new airborne SAR systems and the availability of fully-polarimetric data, new methods can be developed. This paper describes a method to detect contours on fully-polarimetric single-look complex images. The method is based on the multi-variate statistics of speckle in uniform parts of this kind of images. The detector treats the values of the three polarisations as a single data vector as opposed to detecting the contours in each polarisation separately and fusing the results a posteriori. The next section describes some important statistical properties of (polarimetric) SAR images. Then a pre-processor is described that uses some of these statistics to eliminate homogeneous regions from the region to be explored by the contour detector. The fourth section describes the actual contour detector. Section five describes a filter to enhance the results of the detector. The method is applied on an L-band polarimetric SAR image and results are discussed in section six.

2. Some statistical properties of SAR images

2.1. Introduction

The primary geophysical quantity determining the appearance of SAR data is the complex radar reflectivity of the imaged surface. This radar reflectivity expresses the fact that, when an electromagnetic wave scatters from a given position of the earth's surface, the physical properties of the terrain cause changes in both the phase and the amplitude of the wave. In distributed targets each resolution cell of the imaging system contains a large number of discrete scatterers. As the wave interacts with the target, each individual scatterer thus contributes a backscattered wave with a phase and amplitude change. The resolution of the SAR system is typically many times the wavelength of the radar, hence, even if all elementary scatterers were identical, they would still produce a very different phase in the incident wave as the waves scattered from them have very different path lengths. As the value of a pixel in the image is determined by the coherent sum of all the scattered waves within a resolution cell, this results in a noise-like characteristic that is called *speckle*. Speckle is found in any image produced by a coherent imaging system (e.g. laser, sonar, ultrasound, radar). The statistical properties of the speckle are very important as they can be used to develop detectors that are well-adapted to SAR images.

Another important aspect to be taken into account when developing detection or segmentation methods is correlation: spatial correlation within one image as well as inter-channel correlation.

These three topics are briefly discussed in the next paragraphs.

2.2. Distribution of speckle in the SLC images

In single-polarisation Single Look Complex (SLC) images the speckle in non-textured uniform regions follows a zero-mean normal distribution and the difference between different uniform regions is characterised by a difference in variance between these distributions [7].

The distribution of the complex fully polarimetric data (i.e. HH, HV and HV values in each pixel) was verified to be multi-variate normally distributed, again with zero mean and with the difference between uniform regions characterised by differences between the co-variance matrices.

2.3. Spatial correlation

The pixel spacing in SAR systems is always smaller than the resolution cell. This leads to correlation between neighbouring pixels in the SAR image. Furthermore, as the SAR processing in range and azimuth direction is different, the spatial correlation function is not circular symmetrical as is often the case in images created by optical sensors (visible or infrared). The spatial correlation influences the validity of models used to construct detection or segmentation algorithms. The problem is often neglected. Some authors have incorporated this into the model, leading to very complex methods. This is necessary for low-resolution images. However, when dealing with high-resolution images, the problem can be easily circumvented by sub-sampling the image as long as the number of remaining independent samples is still sufficient for the used segmentation or edge detection method [7].

2.4. Correlation between the different polarisations

In many articles the correlation between co- and cross-polar polarisations (i.e. HH/HV and VV/HV) is supposed to be zero. This is only the case for objects with azimuthal symmetry. It is thus true in most vegetated areas but not in man-made objects (villages). On the images we received, i.e. L-band images from a region around an airfield, we determined the inter-channel (between the different polarisations) correlation for several types of land-cover as well as for the runway of the airfield. For each type of land-cover the average over a number of examples was taken. Results are presented in table 1.

Type		HH	HV	VV
Village	HH	1.0	0.362	0.809
	HV		1.0	0.389
	VV			1.0
Forest	HH	1.0	0.136	0.186
	HV		1.0	0.136
	VV			1.0
Runway	HH	1.0	0.057	0.389
	HV		1.0	0.129
	VV			1.0
Grass	HH	1.0	0.0449	0.577
	HV		1.0	0.0448
	VV			1.0
Fields	HH	1.0	0.032	0.588
	HV		1.0	0.0144
	VV			1.0

Table 1. Correlation coefficients between polarimetric components for various types of land-cover

3. The pre-processing step

The distribution of speckle in uniform regions of the polarimetric SLC images should follow a zero-mean multivariate normal distribution. This can be used to locate homogeneous regions in the SAR image, thus limiting the search space for the contour detector. If samples are multivariate normally distributed, the mahalanobis distance from the samples to the sample mean should follow a χ^2 distribution with degrees of freedom equal to the number of variables [6]. This can be checked using a χ^2-test. We used the test with a 10 % significance level instead of the usual 5 or 1 % levels as we wish to avoid tagging a non-homogeneous region as being homogeneous.

4. Description of the contour detector

The basic principle is often encountered in literature: the image is scanned by a set of two adjacent rectangular windows that are rotated around the current pixel (see fig 1). At each positions and for some discrete orientations a statistical measurement is made in both rectangles. The difference between the statistics in both rectangles is used to estimate the probability that the edge between the two rectangles actually corresponds to an edge in the image.

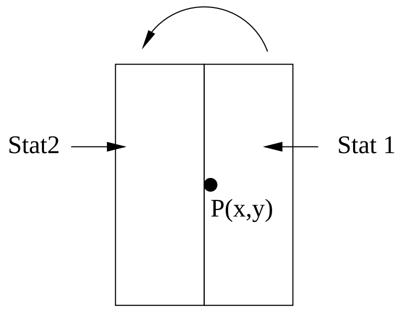

Figure 1. Principle of contour detector

The method focuses on detecting the main contours in high-resolution images, delimiting large uniform regions such as large fields and also the main roads. The detected contours can then be used for registering SAR images with other images or with maps on which the contours of the same objects can be detected.

We have chosen to work directly on the single-look complex (SLC) image. As mentioned earlier, the probability distribution in polarimetric SLC images is zero-mean multivariate normal in uniform regions.

The contour detector problem can then be transformed into a multi-variate hypothesis test, the null-hypothesis being that the covariance matrices are the same in the two rectangular windows. The statistical test that was used to measure the probability that two covariance matrices differ is the Levene test [6]. Using a multi-variate Levene test provides several advantages as opposed to using the uni-variate test for each polarisation and combining (fusing) the results afterwards:

- Different speckle patterns in the different polarisations cause random jitter in the location of the edges. A posteriori fusion will thus result in thickened edges.
- If detectors are run on each channel separately, each detector will have a probability of false targets corresponding to the significance level used in the detector (type one error). Combining these detections will thus combine these errors. If the detector is multi-variate there will only be one component in the error.
- The multi-variate Levene test compares covariance matrices instead of just comparing variances. It thus implicitly takes into account any inter-channel correlation.

The samples from the two adjacent windows are transformed in absolute deviations of sample means. In the case of a single-look complex polarimetric image with complex data of the type x^{HH}, x^{HV}, x^{VV} this results in:

$$\mathbf{L}_{ik} = \begin{bmatrix} \mid Re(x_{ik}^{HH} - \overline{x}_k^{HH}) \mid \\ \mid Im(x_{ik}^{HH} - \overline{x}_k^{HH}) \mid \\ \mid Re(x_{ik}^{HV} - \overline{x}_k^{HV}) \mid \\ \mid Im(x_{ik}^{HV} - \overline{x}_k^{HV}) \mid \\ \mid Re(x_{ik}^{VV} - \overline{x}_k^{VV}) \mid \\ \mid Im(x_{ik}^{VV} - \overline{x}_k^{VV}) \mid \end{bmatrix} \tag{1}$$

in which i is the index of the sample and k the index of the rectangular window. The question whether two samples display significantly different amounts of variance is then transformed into a question of whether the transformed values show a significantly different mean [6]. This can then be tested using a Hotellings T^2 test [1].

The Hotellings T^2 statistic is defined as:

$$T^2 = \frac{n_1 n_2 (\overline{\mathbf{L_1}} - \overline{\mathbf{L_2}})^t \mathbf{C}^{-1} (\overline{\mathbf{L_1}} - \overline{\mathbf{L_2}})}{n_1 + n_2} \tag{2}$$

where n_1 and n_2 are the number of samples used in both rectangles, $\overline{\mathbf{L_k}}$ is the average $\mathbf{L}$ vector in window k and $\mathbf{C}$ is the pooled covariance matrix defined as:

$$\mathbf{C} = \frac{(n_1 - 1)\mathbf{C_1} + (n_2 - 1)\mathbf{C_2}}{n_1 + n_2 - 2} \tag{3}$$

The significance of T^2 is determined by using the fact that in the null-hypothesis of equal population averages the transformed statistic:

$$F = \frac{(n_1 + n_2 - p - 1)T^2}{(n_1 + n_2 - 2)p} \tag{4}$$

follows an F distribution with degrees of freedom p and $(n_1 + n_2 - p - 1)$, where p is the number of variables (i.e. 6 in our case). If we calculate F from the data, we can thus directly associate the probability that the two variances are different to it. However, if the F-statistic (4) is calculated using all pixels in the rectangular windows, the experimentally found 5 % threshold, for a uniform region, is higher than the theoretically predicted one. This can be either due to an intrinsic property of the images and in particular to the spatial correlation or to a problem with Levene's method. If the effect is due to spatial correlation, it should diminish if a smaller percentage of the points in the window are used to calculate the statistic. We thus need to sub-sample. This can be done either with or without replacement. The first step however is to rule out the possibility of a problem with the multi-variate Levene method. In order to do that, we have constructed a set of multi-variate zero mean images without spatial correlation but with inter-channel correlations that are comparable to those actually found in one of the inter uniform regions in a SAR image. The measured F-statistic should have the correct distribution in this simulated image.

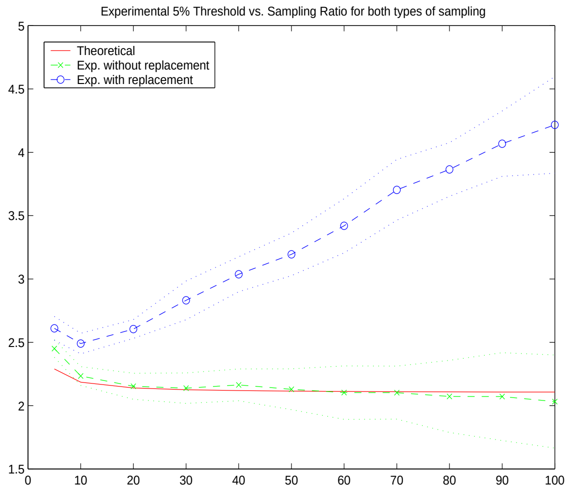

Figure 2. 5% Threshold vs. sampling ratio for sampling with and without replacement for the simulated image

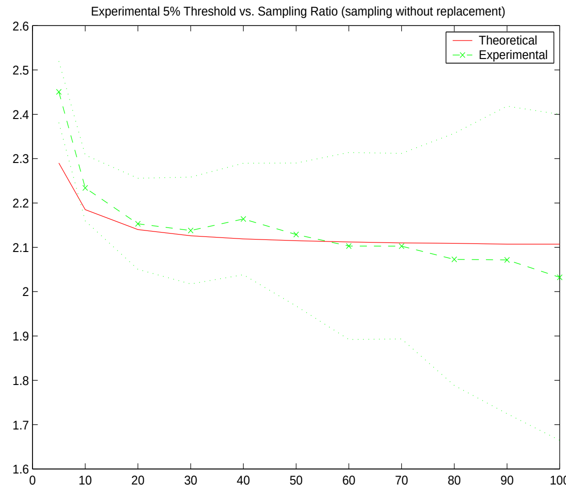

Figure 3. Theoretical and experimental 5% threshold vs. sampling ratio

For this test image we have calculated the 5 % threshold as a function of sampling ratio for sampling with and without replacement. The results are shown in figure 2. When less than 10 % of the points are considered, both sampling schemes are nearly equivalent and their values correspond to the theoretical threshold. However, as the number of selected points increases, the threshold for sampling with replacement increases while the one for sampling without replacement slightly decreases. In fact the behaviour for the sampling without replacement corresponds to the theoretical behaviour of the threshold as shown in figure 3. The fact that in the case of sampling with replacement the threshold increases can be explained by the fact that the replacement itself introduces correlation between the samples (it is possible that the same sample is chosen several times). This experiment thus shows three things:

- The Levene method is valid for our experiment
- Sub-sampling should be done without replacement
- Correlation between samples can explain an increased level for the 5 % threshold

We have thus established that subsampling without replacement reduces the effect of correlation. Now we would like to determine what is the maximum sampling ratio allowed to have results that are useful. Two things need to be checked:

- Do the data indeed follow an F-distribution ?
- Is it possible to assign a constant false alarm threshold independent of the “brightness” of the uniform region?

The answer to these two questions was found by studying the behaviour of the F-statistic in 15 homogeneous regions with different “brightness”. For each region an F-distribution was fitted through the data. A χ^2-test was used to compare the “experimental” F-distribution with the theoretical one. For sampling ratios below 20 % the two distributions are equal (with a significance level of 5 %). At 30 % sampling rate the distributions are only equal if the rectangles are horizontal (as the largest spatial correlation was found to be in the vertical direction, the degradation of the model appears faster using vertical windows). So, sampling should be lower than 30 %.

To answer the second question, the 5 % threshold level was determined experimentally from the histograms of the F-statistic in the 15 uniform regions as a function of sampling rate. The average and standard deviation over the 15 regions was determined. Results are shown in figure 4 for both horizontal and vertical windows. From the figure it appears that it is only possible to assign a global threshold corresponding to a given probability of false targets for all directions when 10 % or less of the points within the window are used to calculate the statistic. Otherwise a threshold has to be set for each direction separately and the method is not valid above 20 % sampling. The detector will thus sample at 10 % and the threshold will be derived from the average experimental F-distribution (see fig 5). As we are only interested in finding contours of large regions, we use a 51×11 rectangular window in the detector. For such windows, a 10 % sampling rate still provides enough independent samples for the method to be valid. As the method will only detect edges between non-textured uniform areas,

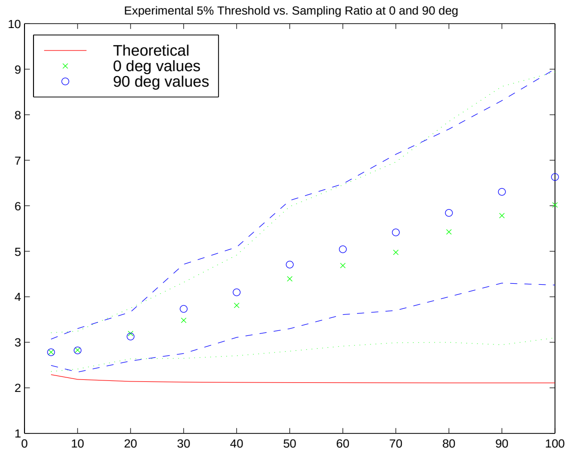

Figure 4. Experimental 5% False Target Threshold vs.sampling ratio for the different (uniform) test regions and for vertical and horizontal windows

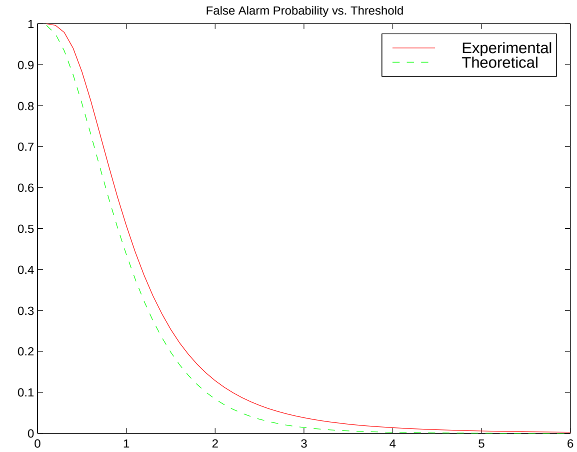

Figure 5. False alarm probability vs. threshold for experimental and theoretical F-distribution

sampling on a fixed grid is equivalent to random sampling (without replacement). This was also checked experimentally and results were indeed found to be equivalent.

5 The post-processing step

In order to filter out individual pixels and to fill small gaps in the detections we used a small morphological operator consisting of an erosion followed by two dilations as a pre-processing step for the binary skeleton filter. The width of the response of the detector depends on the width of the used windows. This results in blurred edges. In order to sharpen these edges, a filter was applied. As the detector is applied using a mask in a given direction, the approximate orientation of candidate edges is known a priori and this information is used in the filter, i.e. a one dimensional filter is applied in a direction perpendicular to the edge. If we know the 5 % false alarm threshold, that threshold can be applied in order to obtain a binary image. The filter finds the centre of binary blocks of given size and replaces its value with the maximal detector output inside that block setting the remainder of the block to zero.

6 Results

The detector was tested on an L-Band SAR image. A part of the HH-polarised component of the original image is shown in fig 6. For the pre-processor, we used 30 by 30 windows with a sampling rate of 20 %. At each step the windows are translated over half the window size. For the contour detector we applied the 1% false alarms threshold, corresponding to a value of 4 for the F-statistic. Results are shown in fig 7. These results were vectorised (see fig 8) using a bar detector to connect the crest of the contour detector's response [5]. As the detector uses a window of 51×11, narrow bars (e.g. narrow roads) are not found. On the other hand most edges between large uniform area are well detected, even if the area contains some texture in which case the assumption of multi-normal distributed speckle is no longer valid.

7 Acknowledgments

The presented work is the result of a collaboration between the Signal & Image Centre of the Belgian Royal Military Academy and the Institut für Hochfrequenztechnik und Radarsysteme of the Deutshes Zentrum für Luft- and Raumfahrt (DLR), who also provided the images. The images used in this project are acquired by DLR's E-SAR system.

8 Conclusions

A multi-variate contour detector for high-resolution polarimetric SAR images was presented. The aim of the detector is to find edges between large uniform regions. We chose to develop a detector for polarimetric SLC images. The detector is based on a multi-variate test of difference in variance between two adjacent rectangular windows located at the current pixel. The obtained measure is then transformed in a probability that the common side belongs to an edge. The rectangles are then rotated around the current

Figure 6. Part of the L-Band image (HH-component)

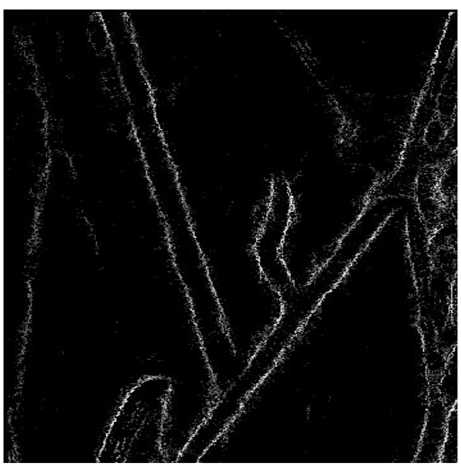

Figure 7. Results of the described method

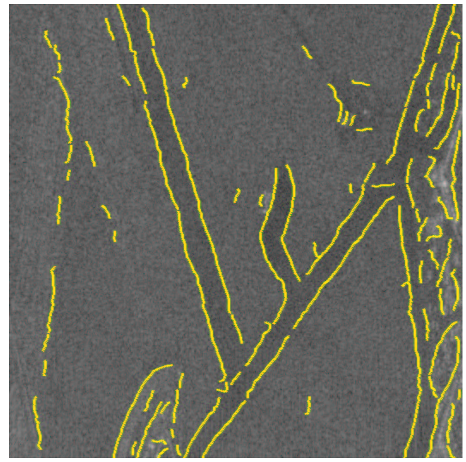

Figure 8. Vectorised contours

pixel. Due to the spatial correlation it is necessary to subsample within these rectangular windows. This ensures that the statistical model is valid such that a constant false alarm rate threshold can be found independent of the “brightness” of the regions. A pre-processor, also based on multi-variate statistics is used to focus the attention of the contour detector to non-homogeneous regions of the image. The developed contour detector is designed to detect only edges between non-textured areas. However, experiments show that edges between a textured (forest or village) and a non-textured area are also detected.

References

[1] T. Anderson. *Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, New York, 1958.

[2] A. Bovik. On detecting edges in speckle imagery. *IEEE-ASSP*, 36(10):1618–1626, October 1988.

[3] J. Canny. A computational approach to edge detection. *IEEE-PAMI*, 8:679–697, 1986.

[4] V. Frost, K. Shanmugan, and J. Holtzman. Edge detection for synthetic aperture radar and other noisy images. In *Proc. IGARSS (Munich)*, pages 4.1–4.9, 1982.

[5] V. Lacroix and M. Acheroy. Feature extraction using the constrained gradient. *ISPRS Journal of Photogrammetry & Remote Sensing*, 53(2):85–94, April 1998.

[6] B. Manly, editor. *Multivariate Statistical Methods*. Chapman and Hall, 1995.

[7] C. Oliver and S. Quegan. *Understanding Synthetic Aperture Radar Images*. Artech House, Boston-London, 1998.

[8] R. Touzi, A. Lopes, and P. Bousquet. A statitical and geometrical edge detector for sar images. *IEEE-GRS*, 26(6):764–773, November 1988.

[9] F. Tupin. *Reconnaissance des Formes et Analyse de Scènes en Imagerie Radar à Ouverture Synthétique*. PhD thesis, ENST, Paris, September 1997.