Chapter 5

Edge Detection in High-Resolution SAR Images

Edge detectors that work well in optical images fail in SAR images. This is due to the particular statistical properties of speckle in SAR images. The most efficient edge detectors for optical images use a gradient-based method, mostly combined with a smoothening filter [27, 28, 29]. These are well adapted for images with additive noise and a high signal-to-noise ratio (SNR). However, the speckle in SAR has the characteristics of a multiplicative noise (in intensity images) with a very low signal-to-noise ratio (SNR). As mentioned in chapter 4 the signal-to-noise ratio for the speckle in a single-look intensity image is equal to one. In multi-look intensity images the SNR is improved by the square root of the number of looks, or more precisely, the square root of the number of equivalent looks (due to spatial correlation) [10, 30]. As multi-look images mostly consist of 4 to 8 looks, the SNR will still be very low. This makes pixel-wise methods using a simple filtering mask inappropriate. Even more elaborate local methods using a combination of an edge detection filter with some smoothening (e.g. [31]) give very noisy results (see fig. 5.1).

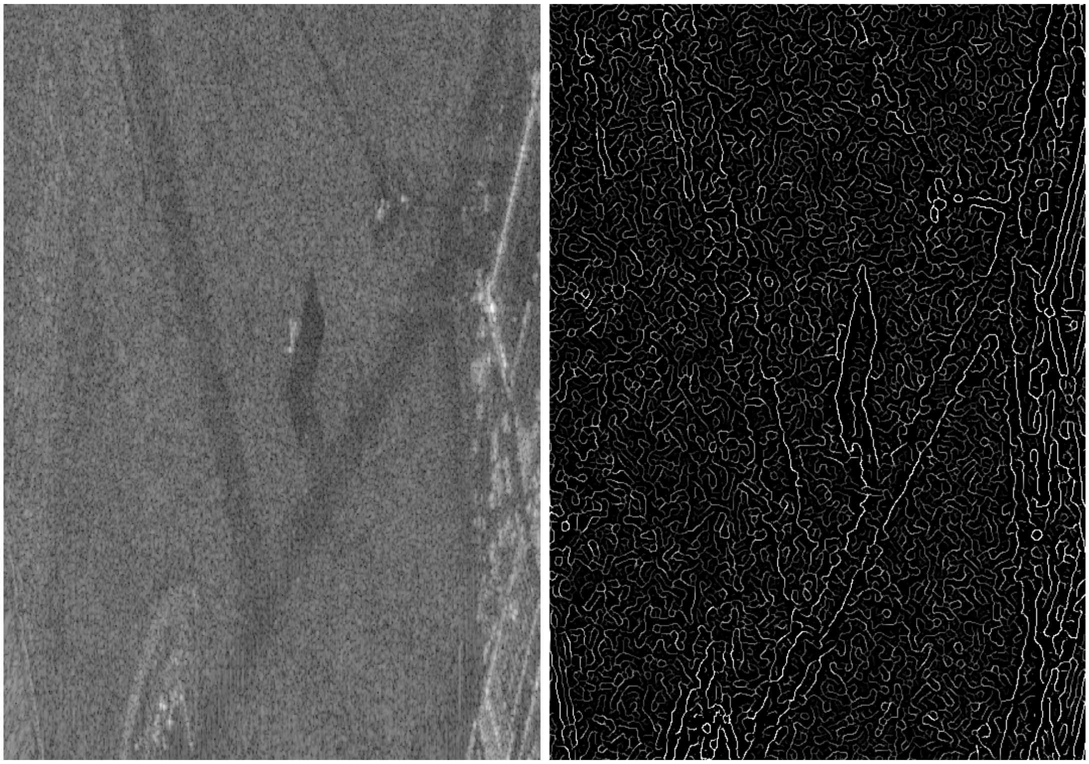


Figure 5.1: Part of HH-polarised log-intensity E-SAR image (left) and result of classical (Canny) edge detector (right)

The solution that is commonly adopted in SAR image processing is to take into account larger neighbourhoods of each pixel for deciding whether an edge passes through that pixel. This is mostly done using the following principle (see fig. 5.2): the images are scanned by a set of two adjacent rectangular windows and in each pixel local measurements, which are a function of statistical differences between the pixels in both windows, are used as an indication that the edge between the two windows actually corresponds to an edge in the image. This gives an answer for one possible edge orientation. The set of rectangles is rotated around the current pixel to verify the presence of an edge along other orientations. Mostly 4 or 8 orientations are tested.

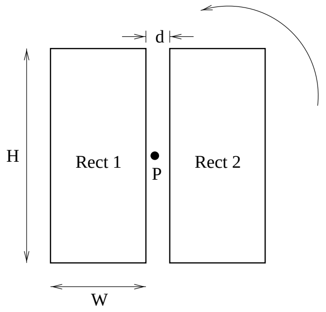


Figure 5.2: General principle for edge detection in SAR images

The existing methods were mostly applied on multi-look intensity images and differ by the comparison criterion that is used. An overview is given in [23] and the most important are enumerated here:

- *Difference in Average Intensity*: Due to the multiplicative character of the speckle, edge detection based on a difference in arithmetic means $d = \overline{I_1} - \overline{I_2}$ between the two rectangles yields a detector of which the probability of false targets increases with the radar reflectivity of the regions [2]. This is a very undesirable property because it means that, in order to obtain a constant false alarm rate (CFAR) detector the threshold used for the detector should be a function of the average intensity around a given pixel. This is very hard to implement. Several solutions are presented in literature. The first is to use the arithmetic average of the log-intensity image. Another possibility is to normalise the difference in means. Two normalisation methods are proposed. The first one normalises by dividing by the sum of average intensities in both windows:

$$d_o = \frac{\mid \overline{I_1} - \overline{I_2} \mid}{\overline{I_1} + \overline{I_2}}. \tag{5.1}$$

  The second one divides by the empirical standard deviation of both windows:

$$d_o = \frac{\overline{I_1} - \overline{I_2}}{\sqrt{\hat{\sigma}_1^2/N_1 + \hat{\sigma}_2^2/N_2}}. \tag{5.2}$$

  These solutions do provide a CFAR detector.

- *Ratio of Average Intensity*: The most widely used criterion is the ratio of average intensities. A detector based on this criterion was proposed by the research teams of Touzi [2] and Bovik [3]. The ratio of average intensities is the maximum likelihood estimation of the ratio of the radar reflectivities in the two rectangles for a single-band SAR image. In that sense the ratio detector is said to be optimal. In order to obtain a result that is bounded and independent of the scanning direction Touzi [2] normalises the average intensity ratio between 0 and 1:

$$r_t = min\left[\frac{\overline{I_1}}{\overline{I_2}}, \frac{\overline{I_2}}{\overline{I_1}}\right]. \tag{5.3}$$

With this definition strong edges give a r_t close to zero. In [30] the detector is therefore defined as $r_t^* = 1 - r_t$.

In [32, 30] the ratio detector has been used for developing a detector for narrow linear features (roads, rivers,...) in single-band SAR images.

- *Hypothesis test*: Oliver [33, 34] proposed a hypothesis test as a log-likelihood difference to compare the two rectangles in single-look intensity images . If both rectangles have the same size the log-likelihood ratio is [34]:

  $$\lambda_D = ln(\overline{I_o}) - ln(\overline{I_1}) - ln(\overline{I_2}), \tag{5.4}$$

  with $\overline{I_o}$ the average intensity in the two rectangles.

The above mentioned method for detecting edges using two adjacent rectangular windows of equal size is sometimes called the *Scanning Window Central Edge (SWCE) method.* All the detectors mentioned above give thick, non-connex edges. The detections need to be post-processed to localise the edges more precisely. This is accomplished by morphological operators or by combining the edge detector with some Gaussian filter. Oliver [34] uses a Fixed Window Scanning Edge (FWSE) approach to localise the position of edges more precisely. Fjortoft [23] combines an edge detector with a watershed method to find closed edges of regions.

In the current chapter we will propose SWCE detectors that exploit the fact that we are dealing with high-resolution SAR images and that we are mainly interested in finding boundaries between large regions that are likely to be present on maps too. Contrary to most methods found in literature, we will be working on single-look images instead of multi-look images. Single-look images provide the highest resolution but the maximum strength of the speckle makes the visual interpretation as well as automatic analysis of such images particularly difficult.

As far as we know there are no publications describing edge detectors that were specifically designed for multi-channel SAR images. The main purpose of the work presented in this chapter is therefore to investigate the detection of edges in (high-resolution) multi-channel (polarimetric) SAR images. An obvious way to detect edges in multi-channel images is to fuse the results of existing detectors applied on each individual channel but we wished to investigate a way to use the information of all channels in the development of the detector. In particular, we investigated the use of multi-variate statistical methods. For multi-channel data we developed two detectors based on multi-variate hypothesis tests. They are presented in section 5.4. We also investigated the use of hypothesis tests for edge detection in single-channel SAR images. Fig. 5.3 illustrates the two approaches that can be used for detecting edges in multi-channel (i.e. polarimetric in our case) SAR images.

In order to speed up the detection process, a pre-processor was developed. It is presented in section 5.2. The pre-processor is based on a multi-variate test for finding non-

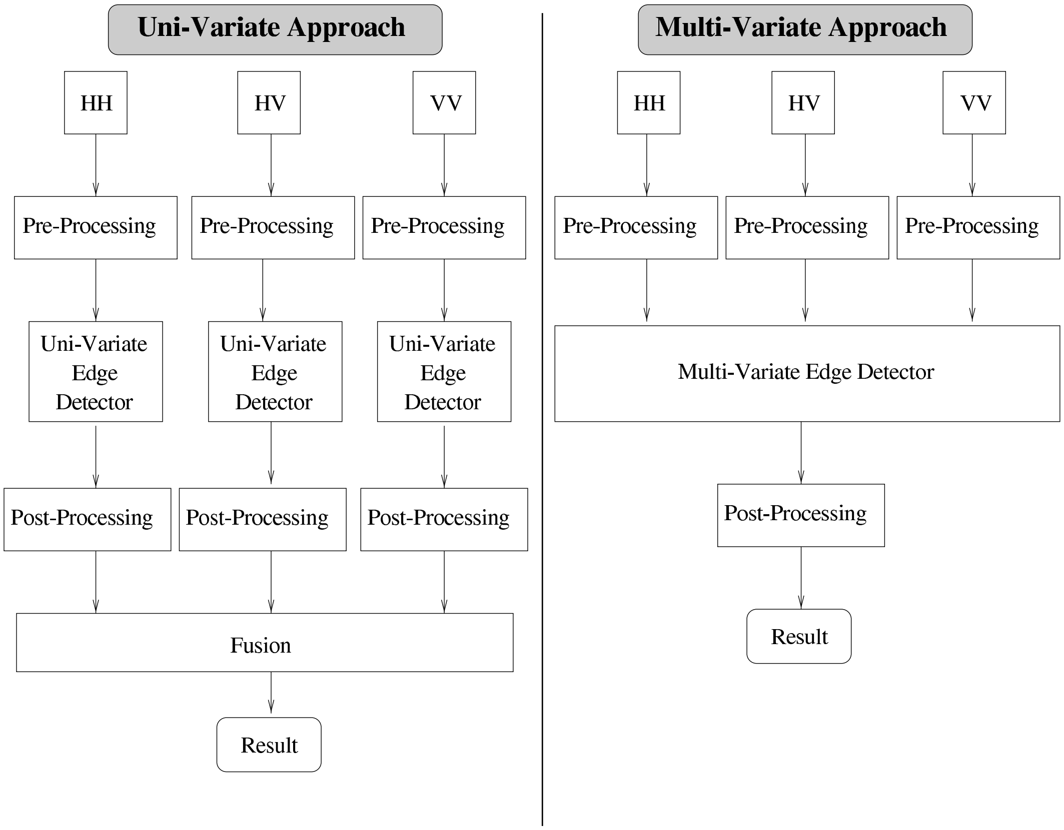


Figure 5.3: Edge detection in polarimetric SAR images

homogeneous regions, thus focusing the attention of the edge detection on "interesting" regions in the image.

The edge detectors discussed in this chapter are all based on statistical hypothesis tests. At each point in the image, in a given direction the presence of an edge is verified using a hypothesis test. The hypothesis H_o is made that the region ($Rect_1$) at the one side of the edge has the same distribution as the region ($Rect_2$) at the other side. The probability of making an error by rejecting H_o is then determined.

Section 5.3 discusses the problem of detecting edges in high-resolution single polarisation SAR images, i.e. it describes the edge detectors based on uni-variate statistics. Two new detectors, based on uni-variate hypothesis tests, are presented here. The first detector is based on a test for the difference of means, the second is based on a test for the difference of medians. Both detectors are applied to log-intensity images where differences in radar reflectivity appear as differences in central tendency of the speckle distribution. When these detectors are applied to the three different polarisations, their results can be fused to give a result for the complete polarimetric image.

Section 5.4 discusses the issue of detecting edges in high-resolution polarimetric (or multi-frequency) SAR images. Two multi-variate edge detectors are constructed here, the first based on the difference of variance and suited for Single-Look Complex images, the second one is based on a test for differences in means and is most suited for log-intensity images. The multi-variate detectors treat the 3 polarisations as a single (vector) data set. As the multi-variate detectors take into account the polarimetric covariance matrix they should be able to use the full polarimetric information more effectively than the uni-variate detectors.

The various edge detectors give a result for each orientation of the scanning windows. Section 5.5 investigates how these results can be combined.

Section 5.6 deals with post-processing the detector's outcome for one orientation of the scanning rectangles. The post-processing deals with narrowing the detector's response and improving the localisation of edges.

In section 5.7 a fusion strategy for edge detection is presented. For the uni-variate approach the fusion of the results obtained for the three polarisations is investigated. On the other hand the two multi-variate detectors have been noted to give complementary results. For that reason their fusion is also examined.

Most publications on edge detection in SAR images neglect the problem of spatial correlation and its effect on the developed edge detectors. An exception is the work of R. Fjortoft [23, 35]. The spatial correlation changes the statistical properties of a detector and makes them differ from the theoretical prediction. When dealing with edge detection in multi-look intensity images, the problem can be circumvented by introducing the notion of equivalent number of looks [30]. In [35] a solution to the problem is proposed by applying a spatial whitening filter as a pre-processing step for edge detection in SLC images.

In the current chapter the effect of spatial correlation on the different developed detectors is studied. The methods developed in this chapter focus on the detection of edges between large uniform regions in high-resolution images. It is therefore possible to use another, simpler method to reduce the effect of spatial correlation: sub-sampling within the scanning rectangles. In the description of each detector these issues are discussed.

In order to compare the results of the various developed edge detectors an evaluation method was defined which is described in section 5.1.

Figure 5.4: Part of SAR image used as test region for edge detectors (L-band, polarimetric E-SAR image with HH: red, HV: green, VV: blue)

5.1 Evaluation Method for Edge Detectors

The results of the different edge detectors will be illustrated on the part of the image shown in Fig. 5.4. The image in the figure is part of an L-band log-intensity E-SAR image.

In order to obtain a quantitative idea of the detector's performance *Receiver-Operator Characteristic (ROC)* curves will be used. ROC curves show the probability of detection of a detector versus its probability of false alarms. They are determined by varying the detector's threshold.

The ROC curves shown in this chapter will be based on simulated images and on the test image of fig. 5.4

Intermezzo: The construction of the simulated images
The simulated images were created as follows. An image with spatially correlated speckle was created. Half of the image corresponds to a given radar reflectivity and the other half differs with a given number of dB from that reflectivity. In order to be able to compare the results obtained with the univariate detectors and those obtained with the multi-variate detectors the three generated polarisation images have the same radar reflectivity. The contrast is defined as the ratio of the total intensities, which, expressed in dB yields:

$$\begin{aligned} Contrast(dB) &= 10\, log_{10}\left(\frac{I_{1,HH}+I_{1,HV}+I_{1,VV}}{I_{2,HH}+I_{2,HV}+I_{2,VV}}\right) \\ &= 10\, log_{10}\left(\frac{\sigma^2_{1,HH}+\sigma^2_{1,HV}+\sigma^2_{1,VV}}{\sigma^2_{2,HH}+\sigma^2_{2,HV}+\sigma^2_{2,VV}}\right) \end{aligned} . \tag{5.5}$$

For the generation of the simulated images the inter-channel correlation that was found in a field (HH/HV: 0.045 HH/VV:0.58) was mimicked. Further details on the method used to generate the simulated images are given in annex A.1.

In order to determine the ROC curves we need to find the probability of detection and the probability of false alarms. As the detectors give a relatively wide response to an edge, this is not straightforward. We have chosen the following definitions which are consistent with the definitions used in [23]:

- *Probability of Detection (P_d)*: For a given threshold this is based on the number of pixels in the detector's result image for which the value is above the threshold and the position is within a given distance of the edge. In [23] this is done by counting the pixels within a given band ($D \leq \Delta_{max}$) around the edge. This means that the result of the detector has to be post-processed (skeletonised) and the result will be biased by the quality of this post-processing. We have chosen to follow the edge and at each pixel of the edge determine if a detected pixel is close enough. Once such a pixel is found it is set to zero, thus avoiding to count a given pixel twice. The P_d is then defined as the number of detections along the true edge divided by the number of pixels of the edge

- *Probability of False Alarms (P_{fa})*: This is defined as the number of pixels above a given threshold that are not within a given distance ($D > \Delta_{min}$) from a true edge divided by the total number of pixels in the image that are not near an edge.

The ROC curve is generated by varying the threshold. For P_d in the simulated image a maximal distance of 1 pixel is taken ($\Delta_{max} = 1$) and the border for false alarms was set to the width of the window ($\Delta_{min} = 11$). For real images, where the edges are not perfectly aligned with one of the orientations of the rotating rectangles, other distances have to be chosen.

Whereas the “ground truth” of correct edge positions is precisely known in the simulated images, this is not the case for the test image. The “true” edges were indicated manually on the image. For the ROC curves of the results of the different contour detectors applied on the test image of fig. 5.4 we have used $\Delta_{max} = 10$, $\Delta_{min} = 11$.

5.2 Pre-Processor for Edge Detection

The distribution of speckle in uniform regions of the polarimetric SLC images should follow a zero-mean multi-variate normal distribution. This can be used to locate homogeneous regions in the SAR image, thus limiting the search region for the edge detectors. If samples are multi-variate normally distributed, the square of the Mahalanobis distance from the samples to the mean should follow a χ^2 distribution with degrees of freedom equal to the number of variables [36]. 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.

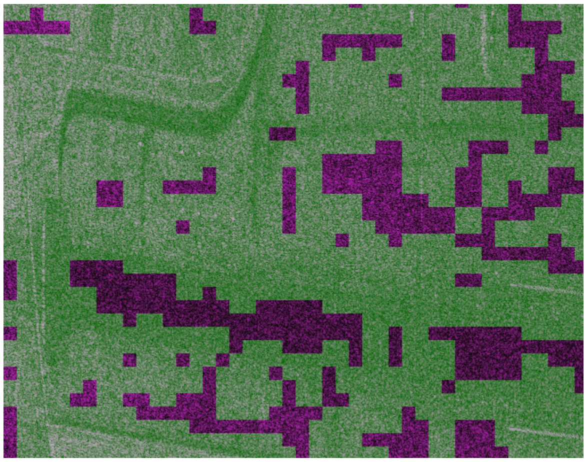


Figure 5.5: Results of the pre-processor superimposed on the image

The pre-processor is able to correctly identify 30 % of the test image as being homogeneous. The only problems are at the level of one pixel wide edges crossing a uniform region (paths or small fences).

5.3 Uni-Variate Edge Detectors

When developing an edge detector for low-resolution images such as those from ERS1, one is dealing with images with a spatial resolution of tens of meters (25m for ERS1). Hence if the edge detection is based on the difference between some statistics on the left and right side of a hypothetical edge, the correct distribution for fully developed speckle in uniform regions has to be used. The average will also follow that distribution.

The fact of having a higher resolution available to detect edges means that larger rectangular windows can be used to detect edges of the same importance. Using larger windows has two consequences. The first is that weaker edges can be detected [2], but, perhaps more importantly: if the number of pixels in the two rectangular windows (assumed uniform) used in the edge detection is large enough, the distribution of their averages approaches a normal distribution, according to the central limit theorem. If this is true, a method based on normal distributions can be used for the edge detection.

That this is indeed the case was experimentally verified in the following manner. We started by delimiting a homogeneous part of the image (about 4000 pixels in size) and then randomly picked 100 pixels and calculated their average. This was done 1000 times and the distribution of the sample averages is examined. The results are shown in fig. 5.6.

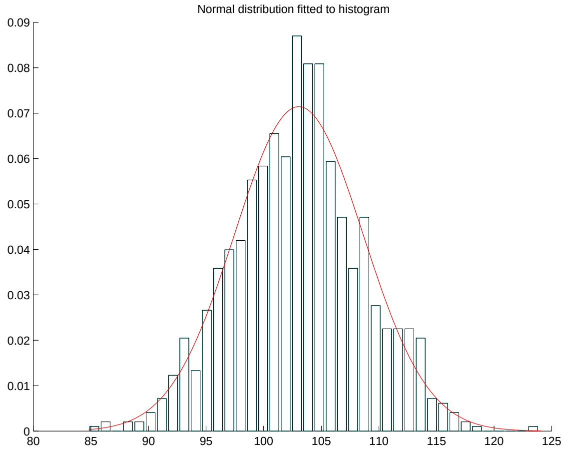


Figure 5.6: Distribution of the average of the intensity image

In the next sections two new contour detectors for high-resolution single-band SAR images are presented. They are both applicable to the log-intensity image. The first detector is based on a hypothesis test for different means. It will be shown that the detector's CFAR behaviour is jeopardized by the spatial correlation in the image unless subsampling is applied. Hence the idea to investigate the use of order statistics for the development of a contour detector. Order statistics could be less sensitive to the problem of correlated observations within a sample. The second detector, presented in sect. 5.3.2, is based on a test for a difference of medians. For all edge detectors we have used scanning windows with a dimension of 11 × 51.

5.3.1 Edge Detection based on a Difference in Means

Principle of the Detector and Test Statistics

The test that is developed here is based on the difference of means. The null hypothesis H_o is that the samples from the two scanning rectangles are drawn from populations with the same mean. The alternative hypothesis H_1 can either be that the mean of one population is different from that of the other or that the mean of one population is higher than the other. In the first case a two-tailed hypothesis test is used, in the second case a one-tailed test is used. As we are concerned here with finding edges we just wish to detect differences in means and thus we will use a two-tailed test.

The average in both regions follows a normal distribution:

$$\begin{array}{l} R_1 \rightarrow \overline{x}_1 \sim \mathcal{N}(E[\overline{x}_1], \sigma_{\overline{x}_1}) \\ R_2 \rightarrow \overline{x}_2 \sim \mathcal{N}(E[\overline{x}_2], \sigma_{\overline{x}_2}) \end{array} \tag{5.6}$$

Note that $E[\overline{x}_i] = E[x_i]$.

The difference of the averages also follows a normal distribution with:

$$\overline{x}_1 - \overline{x}_2 \sim \mathcal{N}(E[x_1] - E[x_2], \sqrt{\sigma^2_{\overline{x}_1} + \sigma^2_{\overline{x}_2}}), \tag{5.7}$$

which is equivalent to:

$$\frac{\overline{x}_1 - \overline{x}_2 - (E[x_1] - E[x_2])}{\sqrt{\sigma^2_{\overline{x}_1} + \sigma^2_{\overline{x}_2}}} \sim \mathcal{N}(0,1). \tag{5.8}$$

If the observations in the two regions are uncorrelated, the variance of the means $\sigma^2_{\overline{x}_i}$ can be calculated from the population variance in each region $\sigma^2_{x_i}$:

$$\sigma^2_{\overline{x}_i} = \frac{\sigma^2_{x_i}}{n_i} \quad i \in \{1,2\} \ , \tag{5.9}$$

with n_i the sample size in region i. The population variances can be estimated from the samples :

$$S_{x_i} = \sqrt{\frac{\sum_{k=1}^{n_i} (x_{i,k} - \overline{x_i})^2}{n_i - 1}} \tag{5.10}$$

If the null-hypothesis H_0 is valid, i.e. if there is no edge between the two regions Rect1 and Rect2, their theoretical averages is the same and we can define a statistic D_{12} which follows a normal distribution:

$$D_{12} = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{\sigma^2_{x_1}}{n_1} + \frac{\sigma^2_{x_2}}{n_2}}} \sim \mathcal{N}(0,1) \tag{5.11}$$

When the population variances are not known but are estimated from the data the difference D_{12} follows a Student t-distribution but if the sample size in both rectangles is sufficiently large (> 30), the Student distribution can be approximated by an $\mathcal{N}(0,1)$ distribution [37].

The obtained test statistic is identical to the test-statistic obtained in the Behrens-Fisher problem which is a variation of the Student t-test for significantly different means with the special conditions that the population variances are not known, i.e. they have to be estimated from the data and that they are not necessarily the same [37]. As the obtained result is a special case of the Student t-test we will refer to the contour detector discussed in this paragraph as the *Student test based contour detector.*

The value d_{12} that is actually found in the image for the test-statistic D_{12} (i.e. d_{12} is a realisation of the test-statistic D_{12}) can be used as a measure of the error that we make by rejecting H_o. This is illustrated by fig. 5.7. In the figure a zero-mean normal distribution with a standard deviation of 1 is shown. For the case where the experimental difference of means is 2 the integral in $[-2, 2]$ is plotted. The error is the sum of the two tails on both sides of the plotted area thus:

$$Error = \frac{1}{\sigma\sqrt{2\pi}}\left[\int_{-\infty}^{-|d_{12}|} e^{-\frac{t^2}{2\sigma^2}}\,dt + \int_{|d_{12}|}^{+\infty} e^{-\frac{t^2}{2\sigma^2}}\,dt\right]. \tag{5.12}$$

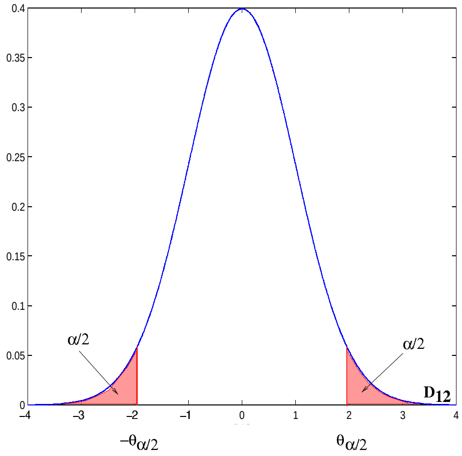


Figure 5.7: Two-tailed error (shaded area in the figure)

The error expressed above is the two-tailed probability of finding an even less probable value for the given distribution , i.e. the p-value. Let $F_{D_{12}}$ be the cumulative distribution function of the test-statistic (i.e. $F_{D_{12}} = F_{\mathcal{N}(0,1)}$ if H_o is valid). The p-value is then defined as:

$$\begin{array}{lll} 2 * F_{D_{12}}(d_{12}) & if & F_{D_{12}}(d_{12}) < 0.5 \\ 2 * (1 - F_{D_{12}}(d_{12})) & if & F_{D_{12}}(d_{12}) \geq 0.5 \end{array} \tag{5.13}$$

If the purpose is only to detect edges the sign of D_{12} is not important and we can replace D_{12} by its absolute value.

The theoretical α % false alarm threshold $\theta_{\alpha/2}$ is then defined as:

$$P\left\{\mid D_{12} \mid > \theta_{\alpha/2} \mid H_o\right\} = \alpha. \tag{5.14}$$

This means that, if the null-hypothesis is rejected for all outcomes of the detector $\mid d_{12} \mid > \theta_{\alpha/2}$, the probability that this rejection is erroneous, is α %.

Influence of Spatial Correlation

The fact that D_{12} defined in eq. 5.11 follows an $\mathcal{N}(0,1)$ distribution relies on the hypothesis that the observations in both regions are uncorrelated. However, when the observations are (spatially) correlated, equation 5.9 has to be adapted; the variance of the mean is no longer equal to the variance of the samples divided by the number of observations. In annex C.1 a correction factor for the variance of the mean as a function of the spatial correlation is derived. The result, which is also cited in [10] is:

$$\sigma_{\overline{x}}^2 = \frac{\sigma_x^2}{WH} \sum_{|i|<W} \sum_{|j|<H} \left[1 - \frac{|\, i \,|}{W}\right] \left[1 - \frac{|\, j \,|}{H}\right] \rho(i)\rho(j), \tag{5.15}$$

where $\rho(i)$ (resp. $\rho(j)$) is the correlation between pixels that are i (resp. j) positions apart along the x (resp. y) axis in the image and W and H are respectively the width and height of the averaging window (the scanning window used in the detector).

If we were to calculate the average in the scanning windows, equation 5.9 will thus underestimate the variance on the average. This means that the actual variance will be higher than the model predicts and that we will classify a given pixel too easily as not belonging to the model distribution. This is what we notice when estimating the test-statistic D_{12} in uniform regions. In such regions the model without correlation predicts that 95 % of the observations should lie in the interval $[-1.96, 1.96]$. However we find values that vary over a much wider range. We thus need to take the spatial correlation into account in the detector. This can be done in several ways:

- A first possibility to take into account the spatial correlation is to replace eq. 5.9 by eq. 5.15, i.e. use the correction factor for the variance of the means. In annex C.1 the correction factor is calculated for 51×11 scanning windows using the average spatial autocorrelation function found in section 4.2.1. For vertical scanning windows the correction factor for a log-intensity image is found to be 5.03 and for horizontal windows it is 4.78. Dividing D_{12} by the square root of this correction factor should thus result in a test statistic that does follow an $\mathcal{N}(0,1)$ distribution. However we will show that this approach is only possible in regions without any texture.

- Another possible solution is to reduce the effect of spatial correlation by spatial sub-sampling in the scanning windows. D_{12} will then be estimated using only a sub-set of the pixels within the scanning windows. As we are interested in detecting edges between large uniform regions in the SAR image, we can use large scanning windows and the sub-sampling will thus not necessarily jeopardize the assumption of normally distributed averages. However, care should be taken how to implement this sub-sampling. We will show that if random sub-sampling is used, it should be without replacement, i.e. making sure that every pixel can only be selected once. Sampling on a fixed grid is an alternative to random sub-sampling and is also examined.

In the next sections we will investigate the behaviour of the detector in several circumstances:

- first its application in simulated images without spatial correlation is considered. This study will show the validity of the detector for the uncorrelated speckle case and it will also allow us to investigate the difference between random sampling with and without replacement;
- next the detector is applied to uniform regions of the actual SAR image, thus in the presence of correlated speckle. The influence of correlated speckle is investigated and a strategy, based on random sub-sampling, is defined to deal with this correlation;
- the third part considers the sub-sampling on a fixed grid
- in the last part some results are discussed.

For this first detector each of these four parts will be discussed in detail. In the discussion of the other detectors the same types of investigations will be presented but their discussion will be kept more briefly.

Application on simulated images without correlation

To verify the validity of the method for non-correlated data we have constructed a set of simulated images.

Intermezzo: The construction of the simulated images
The simulated images are based on the 15 uniform regions that were used for determining the autocorrelation function in section 4.2.1. For the uni-variate detectors, only the HH-component was used for creating the simulated images. Two images with a zero-mean normal distribution are created to simulate the real and imaginary component of the SLC images. From these 2 images all other types of images were derived. In particular the log-intensity image was determined for the detector based on the Student-T test. The 15 simulated normally distributed images have the following properties:

- zero mean
- variance equal to the variance of the HH-component image in one of the 15 uniform regions
- zero correlation between the real and imaginary component
- zero spatial correlation

For these simulated image we have determined the 5 % false alarm threshold as a function of sampling ratio for sampling with and without replacement. The results are shown in figure 5.8. The green full line represents the theoretical value of the threshold. The crosses are the experimentally found values of the threshold for the sampling without replacement and the circles represent the values or sampling with replacement. The represented values are the average over the 15 simulated images. For both sampling schemes the dashed lines represent the average plus and minus the standard deviation of the variation of the threshold over the 15 simulated images. For the sampling without replacement the experimentally determined threshold corresponds to the theoretical one for all sampling rates above 10 %. At 5% sampling rate the experimental threshold increases. This is due to the fact that the resulting sample size is too small (< 30). For the sampling with replacement the threshold is always higher than the theoretical one. The difference increases with increasing sampling rate. 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 observations within a sample (it is possible that the same point is chosen several times). This experiment thus shows that:

- sub-sampling should be done without replacement
- correlation between observations within a sample can indeed explain an increased level for the experimentally found CFAR thresholds

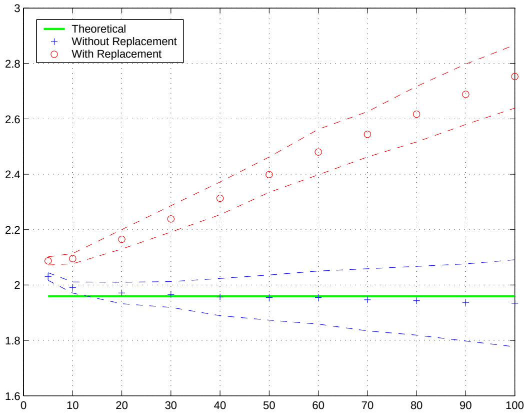


Figure 5.8: 5% False alarm threshold vs. sampling ratio for sampling with and without replacement

Application on uniform parts of the SAR image (with correlation)

For random sub-sampling without replacement the maximal percentage of points that can be selected in order to avoid the problem due to correlation was determined in the following manner. The detector was run in 15 uniform regions of the HH-image and each time the 5 and 1 % false alarm thresholds were experimentally determined. The detector was run by selecting a decreasing number of pixels randomly within the each scanning window to constitute the two samples. This was done for vertically and horizontally oriented windows with a size of 51 × 11. Figure 5.9 shows the thresholds as a function of sampling ratio for vertical and horizontal orientations of the scanning rectangle. The average thresholds for the 15 regions are presented (as symbols) and the average plus and minus one standard deviation (measuring the variation of the threshold over the 15 regions) is also presented (as dashed lines). The theoretical value of the threshold is also shown.

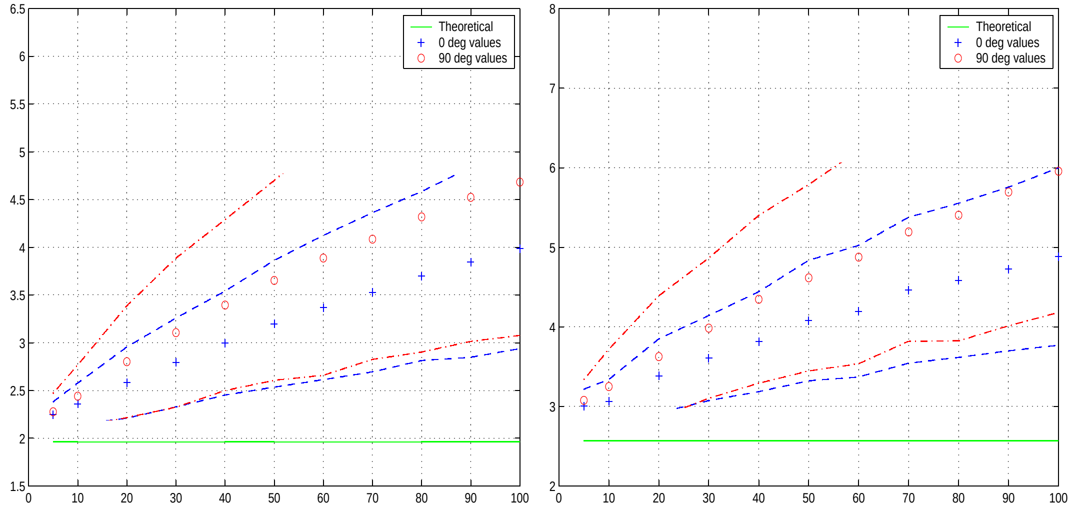


Figure 5.9: 5% (left) and 1% (right) threshold vs. sampling ratio for horizontal and vertical windows for the 15 homogeneous regions

Several conclusions can be drawn from the figure:

- Except for sampling ratios below 10% the thresholds found for vertical windows are larger than for horizontal windows. This is a consequence of the fact that the spatial correlation is larger in the vertical (azimuth direction) than in the horizontal (range) direction and the fact that the scanning windows are rectangles.
- The variance of the threshold over the 15 homogeneous regions increases as more pixels are considered. At 10% subsampling one standard deviation is about 10 % of the value of the threshold while if all pixels are considered, the standard deviation amounts to almost 30 % of the threshold. This means that the detector will give more false alarms in some regions than in others. The detector thus doesn't have a constant false alarm rate.

In annex C the theoretical 5 and 1% false alarm thresholds are determined taking into account the correction factor. The results obtained there are repeated in table 5.1. These results are for 100 % sampling, i.e. all pixels in the scanning rectangles are used.

5 % FA Threshold		1 % FA Threshold	
Hor.	Vert.	Hor.	Vert.
4.17	4.29	5.47	5.63

Table 5.1: Theoretical 5% and 1% false alarm thresholds taking into account spatial correlation for the Student-test based detector

Comparing the values in table 5.1 with the values at 100 % sampling in fig. 5.9, the experimentally found value for a horizontal scanning window seems to be consistent with the theoretical value, while for vertical windows it is higher than the theoretical one.

However, upon closer examination of the 15 "uniform" regions the reason for this has become clear. The 15 regions are the same as those used to determine the autocorrelation function (ACF) in 4.2.1. We have seen that some of these regions have an ACF that is very different from the average (see figs. 4.9 and 4.10). This is due to the presence of some texture of the terrain. When only the 9 regions with a correlation function that is near the average are considered and, in the same way as before, the experimental thresholds are determined as a function of sampling ratio, the results presented in fig. 5.10 are found.

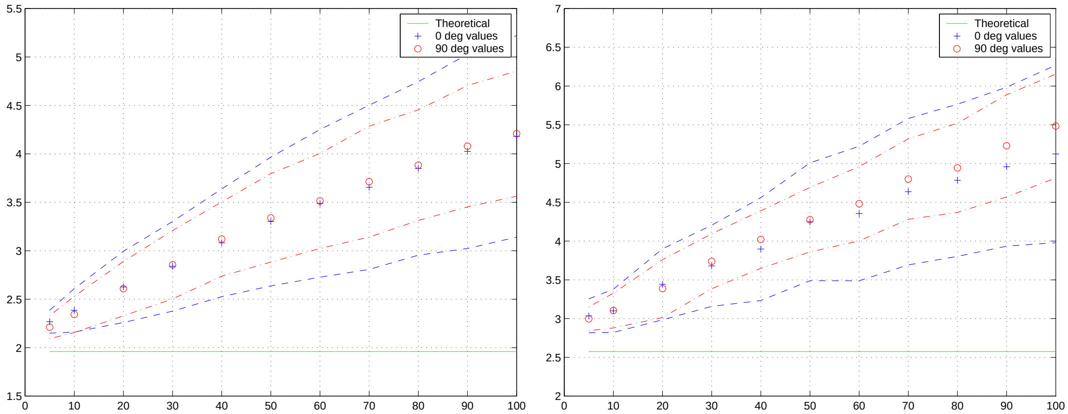


Figure 5.10: 5% (left) and 1% (right) threshold vs. sampling ratio for horizontal and vertical windows for the regions compatible with the ACF

This time the experimental thresholds at 100 % sampling ratio are close to the theoretical thresholds (i.e. the ones corrected for spatial correlation as given in table 5.1) and the variation over the different windows is smaller. This shows that it is only possible to use the detector without subsampling but by applying the appropriate correction factor for the thresholds, in regions were there is no texture. For regions without texture the percentual variation of the threshold over the different regions is independent of the subsampling ratio. In such regions the detector is thus CFAR at all subsampling ratios. Furthermore, in non-textured regions, the variation of the threshold with the orientation of the scanning window is also very small for this detector.

In order to obtain a CFAR detector that is also applicable in (slightly) textured regions it is necessary to subsample. If subsampling is below 20 % the same threshold can be used for the different orientations of the scanning window. Furthermore, at higher subsampling ratios the variation of the threshold over different (textured) regions quickly increases. We will thus sub-sample at 10 %. In highly textured regions the detector will still have a higher false alarm rate.

Intermezzo: Verification of normality condition after sub-sampling
We verified that, if only 10 % of the points in the window are chosen, the normality assumption for the average is still valid. This was done as follows. For a given uniform region a 11×51 window was randomly positioned. From this window 10 % of the points was randomly selected (without replacement) and their average was calculated. This was repeated 200 times. Through the averages a normal distribution was fitted and the difference between the histogram and this normal distribution was verified using a Kolmogorov-Smirnov test. This was done for the 15 uniform test regions. All samples of averages were found to be normally distributed (p-value: $0.6 - 0.9$).

Although by subsampling at 10% a CFAR detector is obtained, from fig. 5.10 it appears that even at such a low sub-sampling ratio, the experimental threshold is still slightly above the theoretical one. In order to find the relation between the probability of false alarms and the detector's threshold, a set of simulated images with spatially correlated speckle was generated. The method used for generating these images is described in annex A.1.

For the 15 uniform regions we simulated images with the same radar reflectivity and inter-channel correlation as the ones found in the corresponding region of the SAR image. A spatial correlation was simulated that is equal to the average spatial correlation found in the uniform regions. By applying the detector (with 10 % subsampling) on these images it is possible to determine the relation between the probability of false alarms P_{fa} and the detector's threshold. The results are shown in fig. 5.11. The red curve is the theoretical behaviour. The solid blue line is the result from the simulated images averaged over the 15 regions and the 8 directions of the detector masks (labeled "Simulated" in the figure). The solid green line (labeled: "experimental") is the result for the 15 uniform regions in the SAR image. The dashed lines represent the average increased and decreased with one standard deviation.

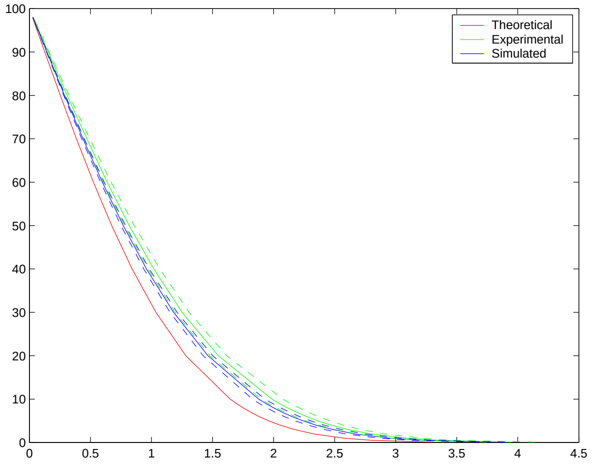


Figure 5.11: Probability of false alarms vs. detector threshold for the Student-test based detector (at 10 % Subsampling)

The variation of the threshold over the different regions and the different directions of the detector's scanning windows is indeed very small. There is a small difference between the simulated curve and the curve determined in the regions of the actual SAR image.

The difference between the simulated curve and the experimental one is due to the fact that some of the selected regions in the real image do contain some texture, increasing their spatial correlation. The curves were determined based on a 10 % sampling ratio.

Sampling on a Fixed Grid

Until now we have reduced the effect of spatial correlation by random subsampling without replacement. An alternative is to sample on a fixed grid. We have defined a grid based on the tables of the autocorrelation function for the log-intensity image (table 4.3). The grid was defined such that the average spatial correlation between the different elements of the grid is smaller than 10 %. This corresponds to a displacement of 3 pixels in azimuth and 1 pixel in range direction. The corresponding grid is shown in fig. 5.12.

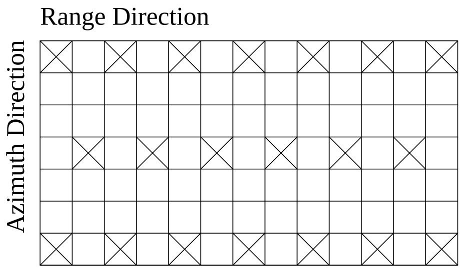


Figure 5.12: Grid for fixed-grid sub-sampling

Sampling on a fixed grid has the advantage that the correction factor can be calculated using the methods described in annex C. In the equations for the various correction factors the autocorrelation function to be used is zero everywhere except for the lags corresponding to the grid points. In this case the equation for the correction with $\rho(i,j)$ needs to be used. The correction factors are presented in table C.3 of the annex. For the Student test based detector the correction factor is 1.02 (the square root of the value in the table).

In fig. 5.13 the probability of false alarms is plotted versus the thresholds for the detector. The shown results are the average over the 15 uniform regions (the circles in the figure) and the corresponding simulated images (the crosses in the figure). The theoretical values that are shown are the theoretical values for the detector multiplied with the appropriate correction factor. The variance over different windows (and directions of detector masks) is very low. For that reason and because the three curves are so close together, the average results from the simulated and real image are shown as discrete points and the standard deviation is not shown.

The sampling on the fixed grid presented in fig. 5.12 results in a subsampling rate of 16.6 %. For 51 × 11 scanning windows 93 points are selected in each scanning rectangle. Sampling on a fixed grid thus allows to have a CFAR detector at higher sampling ratios than when random sampling is used (16.6 % in stead of 10 %). Furthermore it allows to speed-up the contour detection considerably.

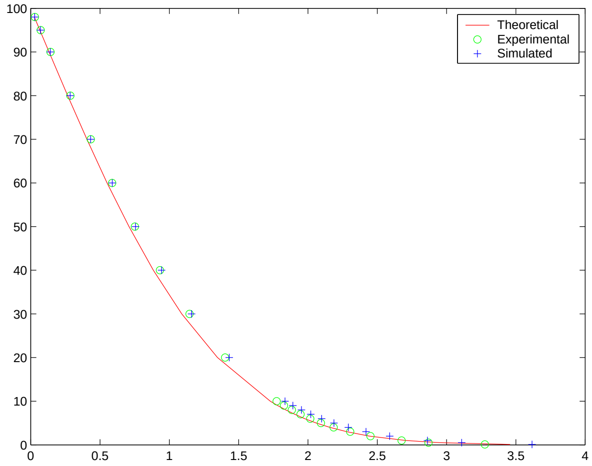


Figure 5.13: P_{fa} vs. Threshold for Student-test based detector for sampling on a fixed grid

Results of the Student Test Based Detector

The raw result of the detector for a given orientation of the scanning rectangles is an image in which each pixel represents the value of the test-statistic, i.e. $\mid D_{12} \mid$ for the Student test based detector. We used 8 orientations of the scanning rectangles. The results of the detector applied on the test image are shown in figs. 5.14 to 5.16. The results shown here are for 10 % random sub-sampling. All images were rescaled such that the interval [0..10] of the detector's outcome was mapped on the greyvalue interval [0..255]. From the figures it appears that the taxiway that runs from the centre bottom of the test image to the top right is only visible on the detector's response for an orientation of the scanning rectangles of 67.5^o (left image in fig. 5.15). This means that the taxiway would not be detected if only 4 orientations of the scanning rectangles would be used. When looking at the figures it is also clear that the detector's response is rather wide; the edges are smeared out. This is not typical for the Student test based detector but, as already mentioned in the introduction to this chapter, it is common to all edge detectors based on a set of scanning rectangles.

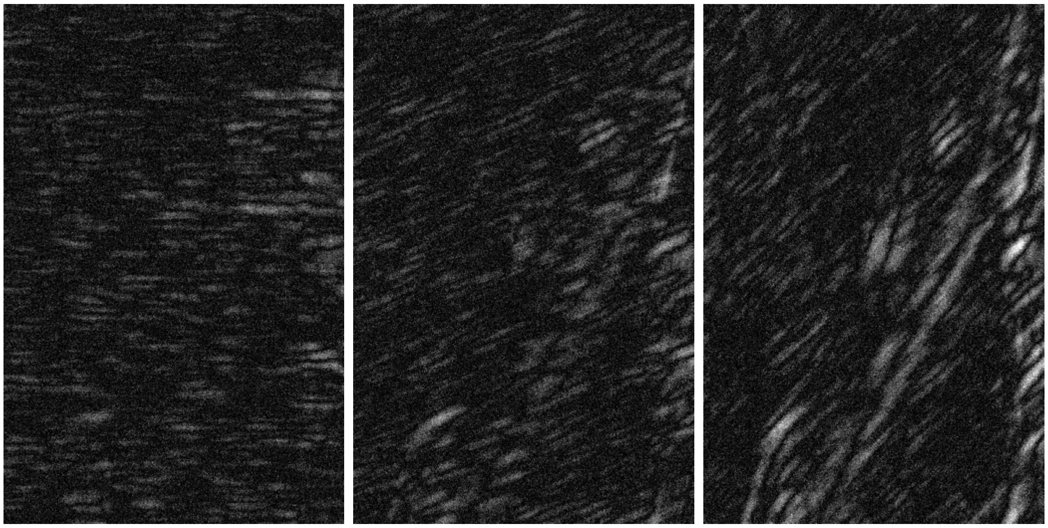


Figure 5.14: Results for 0^o, 22.5^o and 45^o

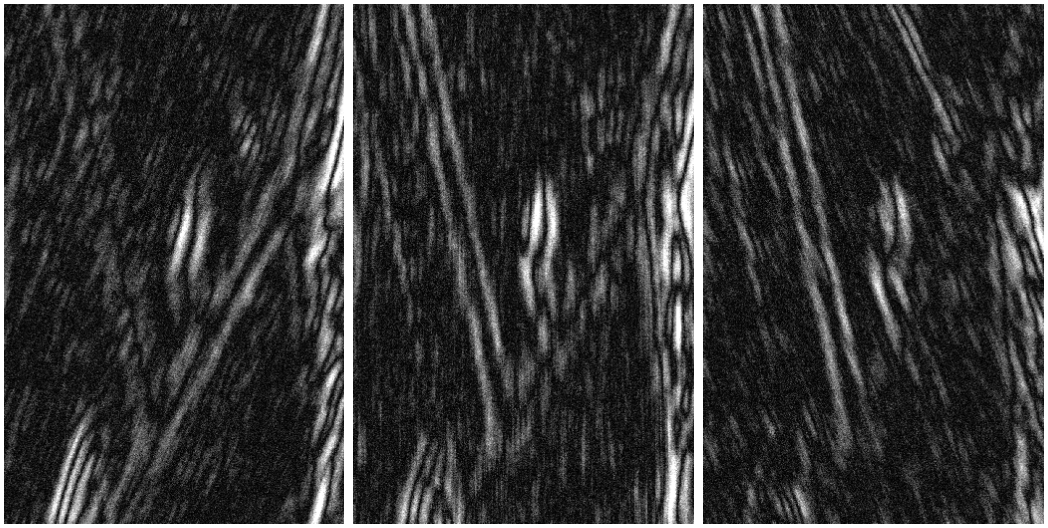


Figure 5.15: Results for 67.5^o, 90^o and 112.5^o

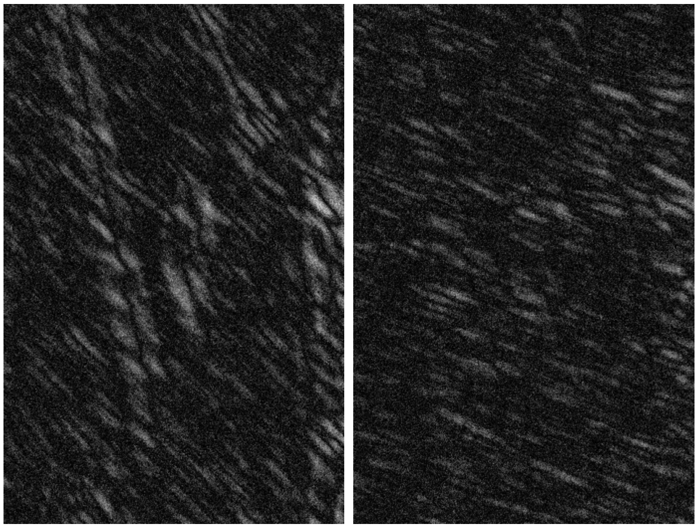


Figure 5.16: Results for 135^o and 157.5^o

The response of the detectors need to be post-processed to improve the localisation of the edges. The discussion of this post-processing is deferred until section 5.6 but we would like to show already the result of such a filtering. In fig 5.17 the result after pre-processing (using the "binary skeleton filter" that will be introduced in sect. 5.6.2) are shown for the different polarisation images. To obtain a single result for a given test image, the responses of the detector for all orientations of the scanning rectangle need to be combined. How this can be done is explained in sect. 5.5. In fig. 5.17 a simple sum was used to combine the response for the different orientations. From the figure it is clear that the various polarisations give different results. The best results for the individual polarisations are obtained in VV but perhaps fusion of the result for the three polarisations can improve the results. This fusion will be investigated in section 5.7.1.

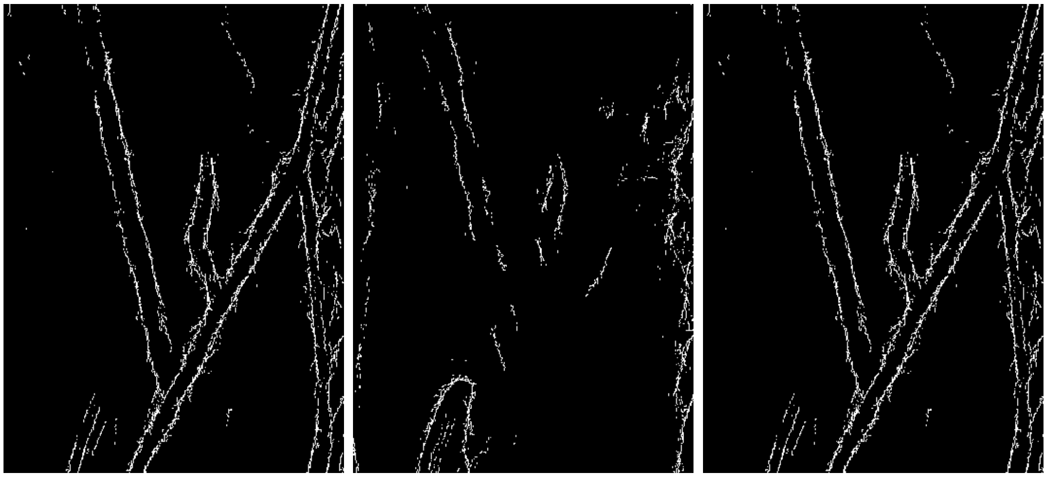


Figure 5.17: Results for HH (left), HV (centre) and VV(right)

In order to compare the different edge detectors, images with correlated speckle were simulated (see sect. 5.1). The resulting ROC curves for different contrast values are shown in fig. 5.18.

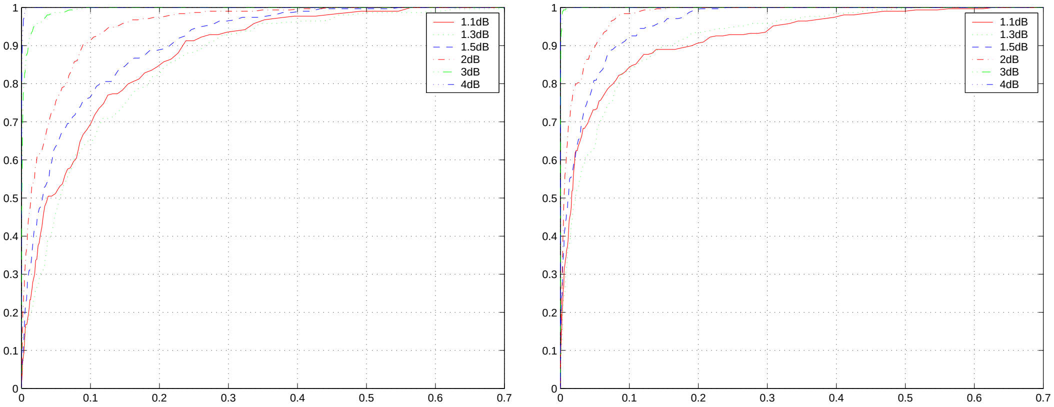


Figure 5.18: ROC for Student-test based detector obtained on simulated images with correlated speckle. Left: Results for random sampling at 10%, Right: Results for sampling on a fixed grid.

The presented ROC curves overestimate the performance of the detector because the detector was run using scanning rectangles with an orientation parallel to the true edge. However these ROC curves are useful because they allow to compare the capabilities of the various edge detectors directly on the raw result of the detector, i.e. without any post-processing and irrespective of any a priori set thresholds. In the left figure the results for 10% random sampling are shown. On the right the results for sampling on a fixed grid are shown. Comparing the values for both sampling schemes at the same contrast level shows that sampling on a fixed grid gives better results. Sampling on a fixed grid reduces the effect of spatial correlation to a low level that is the same in every pixel of the image and the number of points within the sampling window is higher than with random sampling. Another difference between the two sampling schemes is that sampling on a fixed grid results in a grid pattern in the detector's response while the random sampling results in a more erratic response of the detector. This is illustrated in fig. 5.19. Sampling on a fixed grid is thus likely to give better results than the random sampling. However a choice of the best sampling scheme will depend on all the elements of the processing chain. We will therefore continue to use the two sampling methods until the final results of the edge detectors applied on a SAR image can be compared.

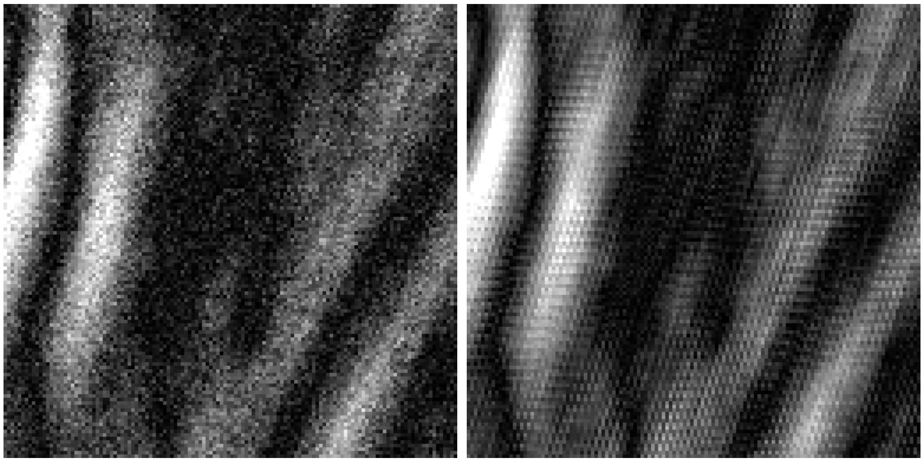


Figure 5.19: Detail of detector's response for random sampling (left) and sampling on a fixed grid (right)

Summary of main conclusions for the Student-test based edge detector

The main conclusions reached by investigation of the Student-test based detector are enumerated below. We will see that these conclusions are also valid for the other developed edge detectors. The most important conclusion is that spatial correlation within the image makes the statistics of the detector differ from the theoretical prediction. The detector thus has to take the spatial correlation into account. This can be done in several ways:

- In regions where the spatial correlation of the speckle is purely due to the PSF of the SAR system (i.e. in non-textured regions), it is possible to use the detector without sub-sampling by introducing a correction factor into the definition of the test-statistic.
- In regions with even a slight texture, the supplementary spatial correlation due to that texture has to be reduced by sub-sampling
  - when random sub-sampling is used it should be without replacement. At 10 % sub-sampling ratio the CFAR thresholds of the detector are independent of the orientation of the scanning windows and of the region itself for the regions of the test image that were investigated. The relation between probability of false alarms and detector's threshold is established using simulated images.
  - an alternative to random sub-sampling is sub-sampling on a fixed grid. For sub-sampling on a fixed grid a sub-sampling ratio of 16.6 % can be used and a correction factor can be applied to find the relation between detector's threshold and probability of false alarms theoretically.

In highly textured regions the false alarm rate of the detector will increase due to the increasing spatial correlation.

5.3.2 Edge Detection based on a Difference in Medians

Principle of the Detector and Test Statistic

We have seen that the spatial correlation jeopardizes the CFAR behaviour of the Student test based contour detector unless subsampling is performed in the scanning windows that are used to compare the statistics. An alternative is to take into account a correction factor but this is only possible for regions without any texture. Tests based on order statistics could be less sensitive to the problem of spatial correlation. We therefore developed a contour detector based on a statistical test for the difference of medians [38]. The first step is to determine the median for the combined set of observations from both rectangular scanning windows (Rect1 and Rect2). Then, using the estimated median, the two sets are dichotomised as follows:

	Rect1	Rect2
Number of observations above median	A	B
Number of observations below median	C	D

Table 5.2: Median test: categorisation of data

If the two rectangles provide samples from a population of which the medians are the same, we would expect that for each rectangle the number of observations below the median is approximately equal to the number above the median. Thus A and B are similar and C and D also. A χ^2 test can be used to check this. If the total sample size is larger than 40 the test statistic [38]

$$T = \frac{N(\mid AD - BC \mid - \frac{N}{2})^2}{(A+B)(C+D)(A+C)(B+D)} \tag{5.16}$$

follows a χ^2 distribution with degrees of freedom $df = 1$. This test is applied to the log-intensity image.

Application on simulated images without correlation

In fig. 5.20 the 5 % false alarm threshold is shown as a function of the sampling ratio for sampling without replacement. The crosses in the figure represent the average threshold over 15 simulated images without spatial correlation. The dashed line represent the average plus and minus one standard deviation of the variation of the threshold over the 15 simulated images. The green line is the theoretical value of the threshold.

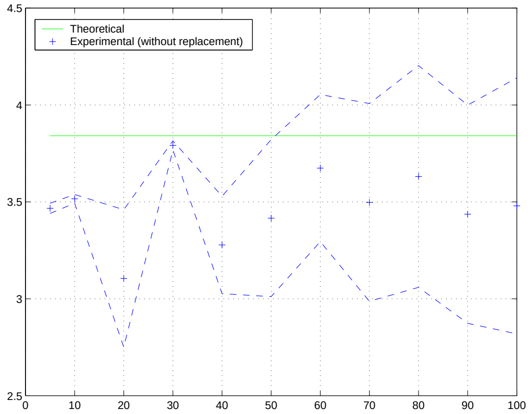


Figure 5.20: 5% (left) False alarm threshold vs. sampling ratio for random sampling without replacement

Application on uniform parts of the SAR image (with correlation)

In order to check the influence of the spatial correlation on the result, the 5 and 1 % false alarm thresholds were determined in homogeneous regions of the Log-Int image as a function of the sub-sampling ratio. As with the previous detector, the thresholds increase when a higher sub-sampling ratio is applied. In fig. 5.21 the results for the 15 homogeneous regions are shown. In fig. 5.22 the results are shown for the 9 homogeneous regions with spatial correlation consistent with 4.2.1. Contrary to what we expected, the median test is even more sensitive to the effect of spatial correlation.

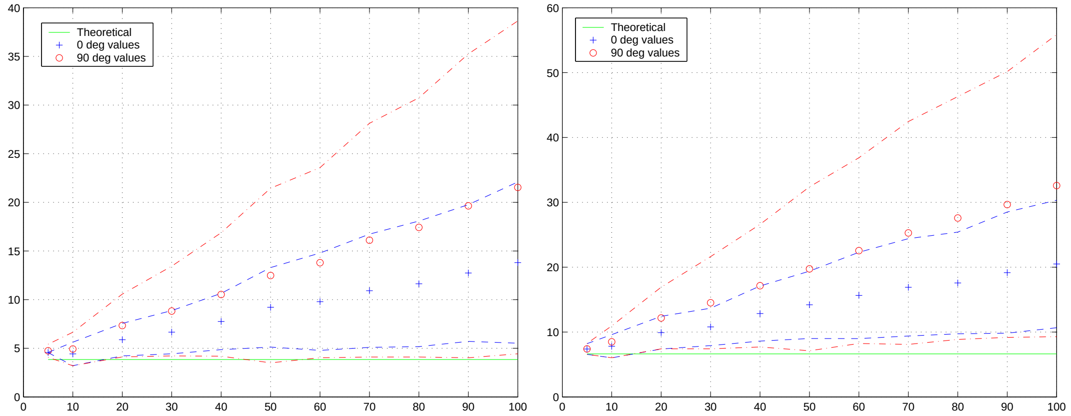


Figure 5.21: 5% (left) and 1% (right) threshold vs. sampling ratio for horizontal and vertical windows for 15 homogeneous regions

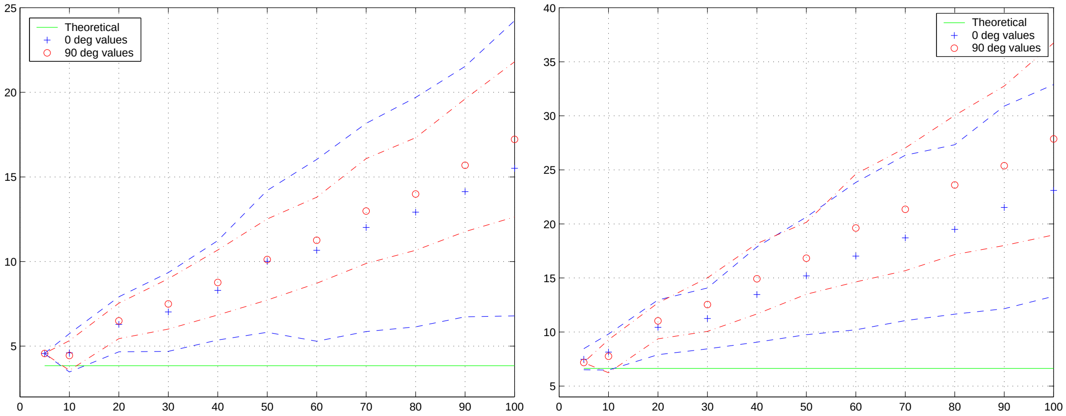


Figure 5.22: 5% (left) and 1% (right) threshold vs. sampling ratio for horizontal and vertical windows for the regions compatible with the ACF

From Fig. 5.21 it seems that we need to subsample at 10 % to obtain a CFAR detector, as in the case of the previous detector. It turns out that we will find this same subsampling ratio for all detectors that were examined. In 5.23 the relation between the detector's threshold and the probability of false alarms is shown. For a detailed description of the meaning of this figure we refer to the text above the corresponding figure for the previous detector (fig. 5.11). For the median detector, when applied with 10% subsampling a good agreement is found between the theoretical curve and the curves found both in the simulated images and the regions of the SAR image.

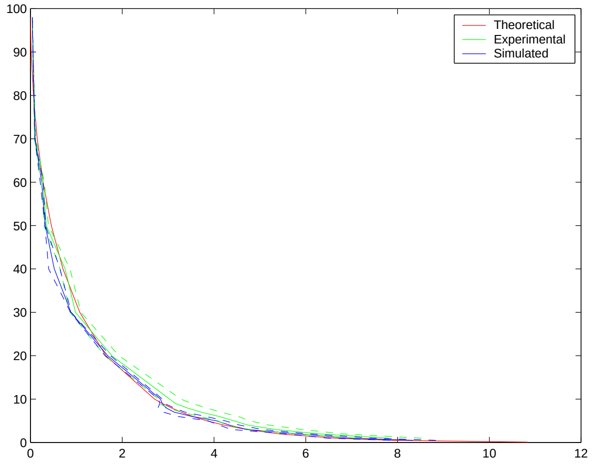


Figure 5.23: Probability of false alarms vs. detector threshold for the median test detector (at 10 % subsampling)

Sampling on a Fixed Grid

In fig. 5.24 the probability of false alarms is represented versus the threshold for sampling on a fixed grid. The same grid is used as for the previous detector and as always the results are shown for 51 × 11 scanning rectangles. No correction factor was applied to the theoretical curve. The results from the simulated images with correlated speckle and those obtained from the uniform regions of the SAR image are both very close to the theoretical curve.

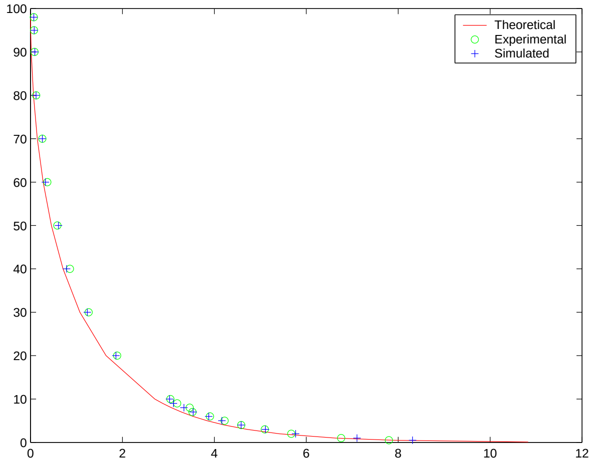


Figure 5.24: P_{fa} vs. threshold for the median test for sampling on a fixed grid

Results of the Median Detector

Fig. 5.25 presents the ROC curves obtained in simulated images with correlated speckle and a known edge orientation. Contrast levels from 1.1 dB to 1.5 dB give very similar results and comparing these results with those obtained for the Student test based detector (fig. 5.18) shows that the Student test based detector clearly outperforms the detector based on the median test.

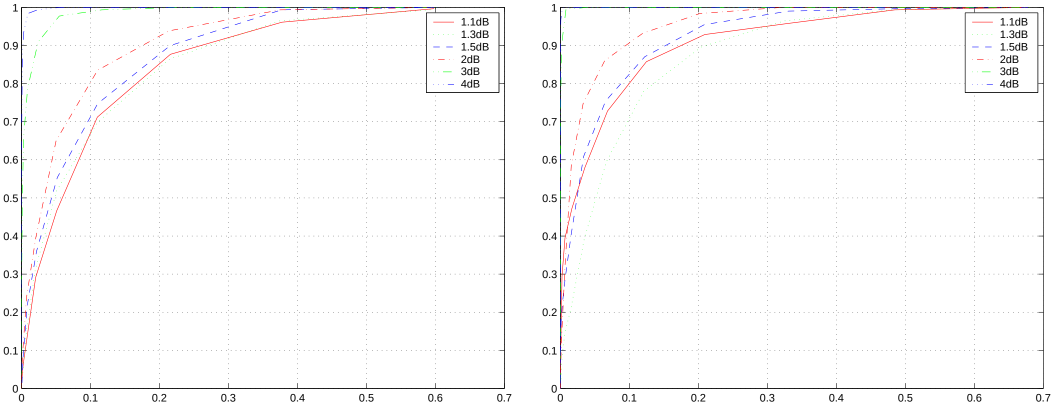


Figure 5.25: ROC for median test based detector obtained on simulated images with correlated speckle. Left: Results for random sampling at 10%, Right: Results for sampling on a fixed grid.

5.4 Multi-Variate Edge Detectors

Generally, when one is dealing with multi-channel segmentation two approaches are possible:

- Uni-variate approach: Segment each channel individually and combine the results.
- Multi-variate approach: Use all channels as one source and segment these combined data.

Both of these methods have possible advantages and disadvantages for edge detection. The first method is often the easiest to implement but, different speckle patterns in the different channels (polarisations) cause random "jitter" in the location of the edges. This results in thickened edges at locations where edge features are visible on all images [10] when the results are combined. Furthermore, the issue of combining the results obtained in the different channels needs to be investigated.

The second method treats the multi-channel image as a single source of edge information, thus avoiding both of these potential difficulties.

Another advantage of the second method is the following. 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. Combining these detections will thus combine these errors. If the detector is multi-variate there will only be one component in the error.

The current section examines multi-variate detectors. Two new multi-variate detectors are proposed.

The first is based on a multi-variate test for the difference in variance between two populations and it is thus suited for single-look complex polarimetric images where differences in radar reflectivity result in a difference of variance of the (zero-mean Gaussian) speckle distribution.

The second detector is based on a test for the difference in means. It was applied to the log-intensity image where differences in radar reflectivity only change first-order moments of the speckle distributon in uniform regions.

5.4.1 Edge detection based on a Difference in Variance

As the real and imaginary parts of the SLC image have a zero-means normal distribution in uniform regions, the vector combining these six measurements should be multi-variate normally distributed. Hence an edge detector could be based on a statistical test for a difference of normal distributions. As the combined vector will have zero mean in homogeneous regions, the test should focus on finding differences between the variances of both regions. In the first part of the section the hypothesis of multi-variate normal distribution is verified. The second part explains the test for difference in variance between two regions. In both parts standard multi-variate statistical methods [36] are used.

Verification of multi-variate Normality

If a distribution is multi-variate normal, the square of the Mahalanobis distances of its samples to the estimated population mean is χ^2 distributed with degrees of freedom equal to the dimension of the multi-variate variable [36]. In order to check this we thus need to identify a uniform region in the SLC images and calculate the mean and covariance matrix of some samples in that region. Then, for each point we calculate the Mahalanobis distance and verify whether this follows a χ^2 distribution. This is done by means of a P-P plot (see fig 5.26).

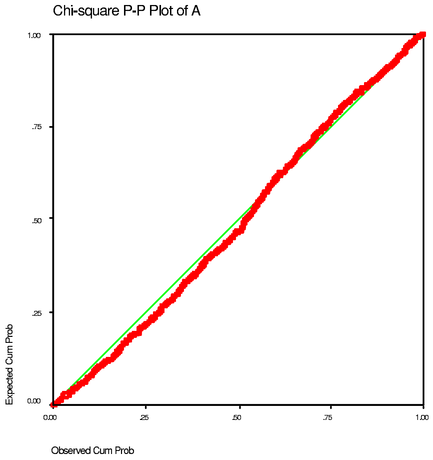


Figure 5.26: Results of P-P plot Mahalanobis distance (χ^2)

Principle of the Detector and Test Statistics

The multi-variate hypothesis test for equality of variances that was used is the Levene test [36]. The null-hypothesis H_o is that the samples from the two scanning rectangles are from populations with the same variance, the alternative hypothesis H_1 is that the population variances are different.

In the Levene test the samples from the two scanning windows are first 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:

$$\mathbf{L}_{ik} = \begin{bmatrix} \mid Re(x_{ik}^{HH} - \overline{x}_k^{HH}) \mid \\ \mid Im(x_{ik}^{HH} - \overline{x}_k^{HH}) \mid \\ \mid Re(x_{ik}^{HV} - \overline{x}_k^{HV}) \mid \\ \mid Im(x_{ik}^{HV} - \overline{x}_k^{HV}) \mid \\ \mid Re(x_{ik}^{VV} - \overline{x}_k^{VV}) \mid \\ \mid Im(x_{ik}^{VV} - \overline{x}_k^{VV}) \mid \end{bmatrix}, \tag{5.17}$$

in which i is the index of the observations and k the index of the scanning window (k=1 or 2). 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. This can then be tested using a Hotellings T^2-test [39, 36].

The Levene test is based on the assumption of normality and equal within-sample variability (covariance matrix) [36]. The normality was verified in the previous section. A difference in population covariance matrices is not too important especially with equal or nearly equal sample sizes [36]. In our case sample sizes are equal and thus we can apply the test.

The *Hotellings T^2-statistic* is defined as:

$$T^2 = \frac{n_1 n_2 (\overline{\mathbf{L_1}} - \overline{\mathbf{L_2}})^t \mathbf{C}^{-1} (\overline{\mathbf{L_1}} - \overline{\mathbf{L_2}})}{n1 + n2}, \tag{5.18}$$

with

$$\overline{\mathbf{L_k}} = \frac{\sum_{i=1}^{n_k} \mathbf{L}_{ik}}{n_k} \tag{5.19}$$

and $[C]$ the pooled covariance matrix estimated by:

$$\mathbf{C} = \frac{(n_1 - 1)\mathbf{C_1} + (n_2 - 1)\mathbf{C_2}}{n_1 + n_2 - 2},, \tag{5.20}$$

where C_1 and C_2 represent the covariance matrices estimated from the two scanning rectangles. The significance of T^2 is determined by using the fact that in the null-hypothesis of equal population means the transformed statistic

$$F = \frac{(n_1 + n_2 - p - 1)T^2}{(n_1 + n_2 - 2)p} \tag{5.21}$$

follows an F_{ν_1,ν_2} distribution with degrees of freedom $\nu_1 = p$ and $\nu_2 = n_1 + n_2 - p - 1$. p is the number of variants, i.e. 6 in our case if the real and imaginary components for each polarisation are counted separately.

From the theoretical distribution of the test statistic the theoretical α% false alarm threshold θ_α for the detector can be determined. It is given by

$$P\{F \geq \theta_\alpha \mid H_o\} = \alpha. \tag{5.22}$$

However the Hotellings test supposes that the covariance of the mean is equal to the covariance of the samples divided by the sample size [39]. As shown in section C.2 of the annex this is only true if the observations are uncorrelated. Spatial correlation introduces the need for a correction factor in a similar way as for the Student test based detector.

Application on simulated images without correlation

In order to verify the validity of the multi-variate detectors for non-correlated observations, we need to simulate polarimetric images without spatial correlation. The HH-polarisation real and imaginary component images of the SLC, created for the simulation of single-band non-correlated images in section 4.2.1 were used as a starting point. The other polarisation images were derived from this image using the method described in annex A.2. This ensures that the created images have inter-channel correlations that are comparable to those actually found in each of the uniform regions in a SAR image. From the set of simulated polarimetric SLC images, the other types of images are derived if needed.

For these simulated image we have determined the 5 % false alarm threshold as a function of sampling ratio for sampling without replacement. This time the theoretical distribution of the test statistic depends on the number of observations (pixels) used in both scanning windows and thus of the sampling ratio. We have determined the theoretical values of the 1% and 5% false alarm thresholds for different sampling ratios in 51 × 11 scanning windows. The test statistic follows an $F(\nu_1, \nu_2) = F(6, n_1 + n_2 - p - 1)$ where $n_1 = n_2 = (51 \times 11) * S_p/100$ with S_p is the sampling ratio in percent. The resulting theoretical thresholds are presented in table 5.3.

Sp	N	ν_2	5% Treshold of $F(6, \nu_2)$	1 % Threshold of $F(6, \nu_2)$
5 %	28	49	2.290	3.195
10 %	57	107	2.185	2.975
20 %	114	221	2.140	2.885
30 %	171	335	2.126	2.856
40 %	228	449	2.119	2.842
50 %	286	565	2.115	2.834
60 %	343	679	2.112	2.829
70 %	400	793	2.110	2.825
80 %	457	907	2.109	2.822
90 %	514	1021	2.107	2.820
100%	572	1137	2.107	2.818

Table 5.3: Theoretical 5% and 1% false alarm threshold for F-distribution vs. sampling percentage in a 51x11 window

The behaviour for the sampling without replacement in simulated images without spatial correlation corresponds to the theoretical behaviour of the threshold as shown in figure 5.27.

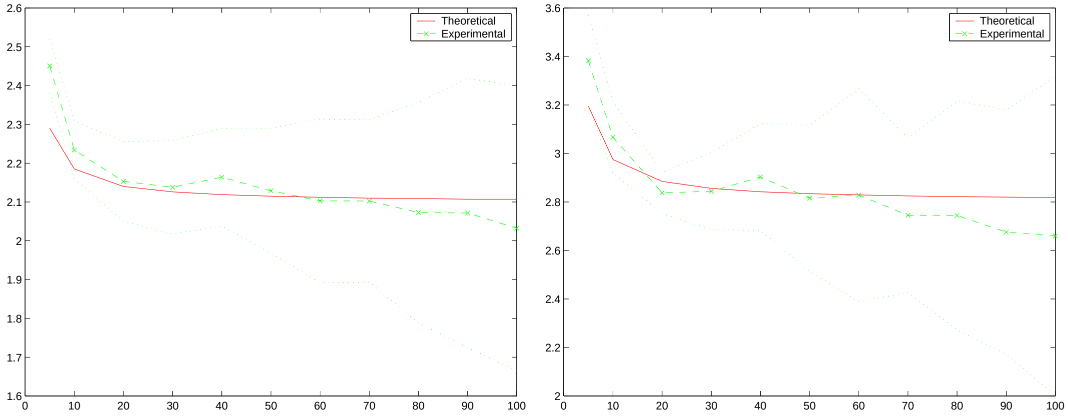


Figure 5.27: 5% (left) and 1% (right) threshold vs. sampling ratio for sampling without replacement and the theoretical threshold

Application on uniform parts of the SAR image (with spatial correlation)

To check the validity of the method in images with correlated speckle the F-statistic was determined in the 15 uniform regions for 8 orientations of the scanning windows and for various sampling ratios. For the resulting F-statistics, the two degrees of freedom for an F-distribution were determined from the data and the goodness of fit between the theoretical and the "measured" F-distribution was determined using a χ^2 test. If data are F-distributed the degrees of freedom of the F-distribution can be estimated from the first and second moment of the experimental distribution [40]:

$$E[F] = \frac{\nu_2}{\nu_2 - 2} \tag{5.23}$$

$$\sigma^2[F] = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2^2 - 2)^2(\nu_1 - 4)}. \tag{5.24}$$

Estimating $E[F]$ and $\sigma^2[F]$ from the data and analytically inverting the equations above yields the two degrees of freedom.

On the other hand the theoretical degrees of freedom are determined from the sampling ratio (cf. Table 5.3). The theoretical and best-fit F-distribution are compared using a χ^2 test. Results show that an F-distribution is found (with a significance level of 0.05) only when less then 30 % of the pixels in the windows are used to calculate the statistics (see figure 5.28). Furthermore, especially above 20 % the quality of the fit with the F-distribution degrades as the longest side of the scanning windows gets more aligned with the azimuth direction, i.e. when the vertical dimension of the scanning window grows. This observation is consistent with the idea that increased spatial correlation degrades the validity of the model.

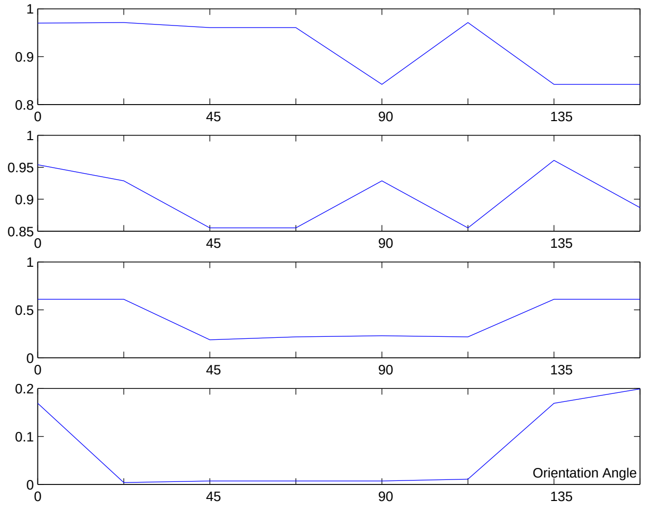


Figure 5.28: Confidence level of χ^2 test vs. orientation of window for different sampling ratios

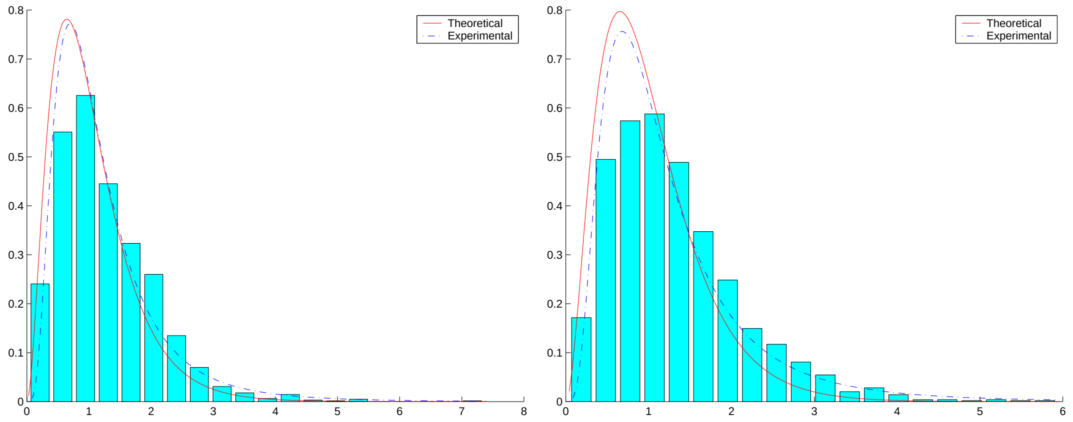


Figure 5.29: Histogram and experimental (blue) and theoretical (red) F-distribution for 5% (left) end 10% (right) sampling ratio

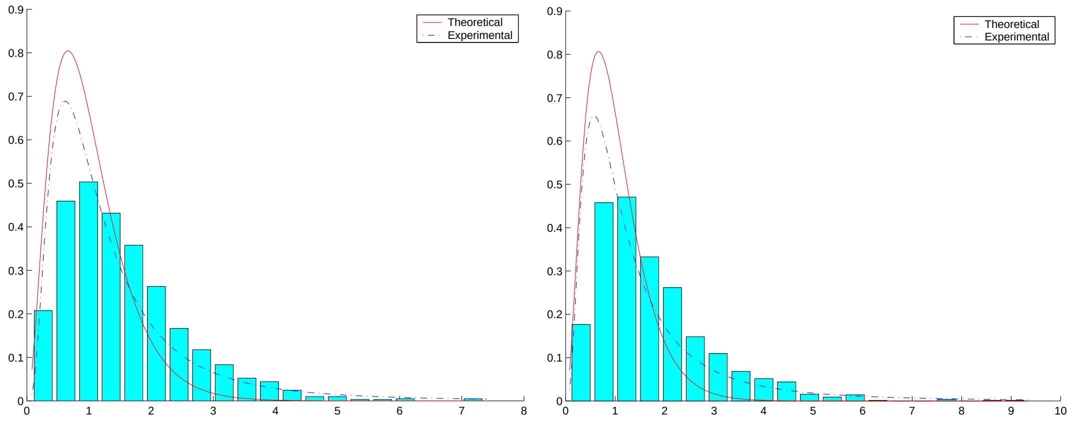


Figure 5.30: Histogram and experimental (blue) and theoretical (red) F-distribution for 20% (left) end 30% (right) sampling ratio

We have also determined the experimental 5 and 1% thresholds as a function of the sampling ratio for 15 uniform regions and for horizontal and vertical directions of the scanning window. Results are shown in figure 5.31. Note that at 5 and 10 % sampling ratio the experimental threshold is almost equal to the theoretical threshold although we find a F(13,11) distribution in stead of a F(6,49) distribution. In fact, although the degrees of freedom are very different, the two distributions are very similar. The significance of dealing with the same distribution, calculated by means of a χ^2 test between F(13,11) and F(6,49) is 0.97. In figure 5.31 the average value over all test images is shown for each value of the sampling ratio as well as the lines corresponding to the average increased and decreased with one the standard deviation (dashed lines). The ratio of the standard deviation of the threshold to its average increases rapidly with increasing sampling ratio. In figure 5.32 the thresholds are shown as a function of sampling ratio for the homogeneous regions that have a spatial correlation that is compatible with the one given in 4.2.1. Here the ratio of standard deviation to average remains constant with the sampling ratio. The thresholds at 100 % sampling ratio are also closer to (but still above) the theoretical thresholds derived in annex C.3 and repeated in table 5.4.

5 % FA Threshold		1 % FA Threshold	
Hor.	Vert.	Hor.	Vert.
5.47	6.44	7.32	8.62

Table 5.4: Theoretical 5% and 1% false alarm thresholds taking into account spatial correlation for the Levene test

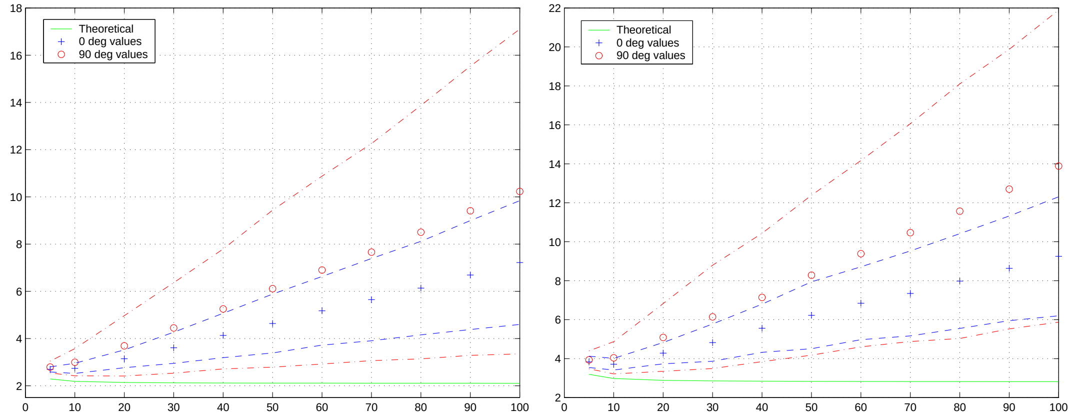


Figure 5.31: Experimental 5% (left) and 1 % (right) false target threshold vs. sampling ratio for the 15 homogeneous regions

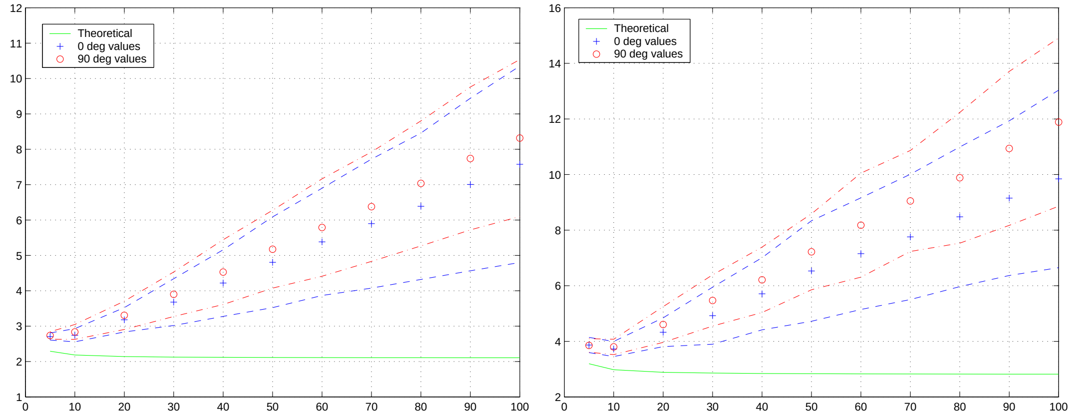


Figure 5.32: Experimental 5% (left) and 1 % (right) false target threshold vs. sampling ratio for the different regions compatible with the ACF

It is thus only possible to assign a global threshold corresponding to a given probability of false target level 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.

Even at 10 % subsampling the experimental threshold is again above the theoretical one. It is thus not possible to use the theoretical F-distribution to find the CFAR threshold. In order to find the thresholds we have determined the threshold as a function of the probability of false targets for the 15 uniform image windows with varying brightness (and for the 8 orientations of the rectangles). In fig. 5.33 the probability of false alarms is plotted (green line) against the threshold averaged over the 15 windows and the 8 orientations. The curves of average minus and plus one standard deviation are also shown as well as the theoretical curve. The variance is very small and thus, for the 15 regions, the detector behaves as a CFAR detector.

For the 15 regions we also created simulated images with the same radar reflectivity and interchannel correlation as in each of the original regions and with the average spatial correlation found in uniform regions. For these simulated windows the probability of false alarms versus the threshold was also determined. The results are shown in blue in fig. 5.33. The dimension of the simulated images is 200×200 which ensures that thresholds corresponding to very low false alarm rates can be determined.

The difference between the simulated curve and the experimental one is very small and we will use the results from the simulated images to determine the CFAR threshold.

The detector will thus subsample at 10 % and the thresholds will be derived from simulated images.

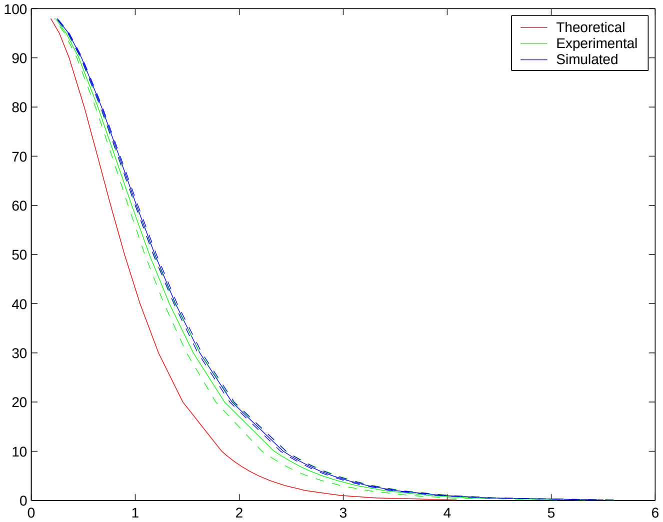


Figure 5.33: Probability of false alarms vs. detector threshold for the Levene test

Sampling on a Fixed Grid

For sampling on a fixed grid the experimentally found thresholds in uniform regions of the SAR image and in simulated images with correlated speckle agree with the theoretical thresholds multiplied with the appropriate correction factor for the single-look complex image presented in table C.3 of the annex. The correction factor used in the figure was 1.14 which is the average value for vertical and horizontal windows.

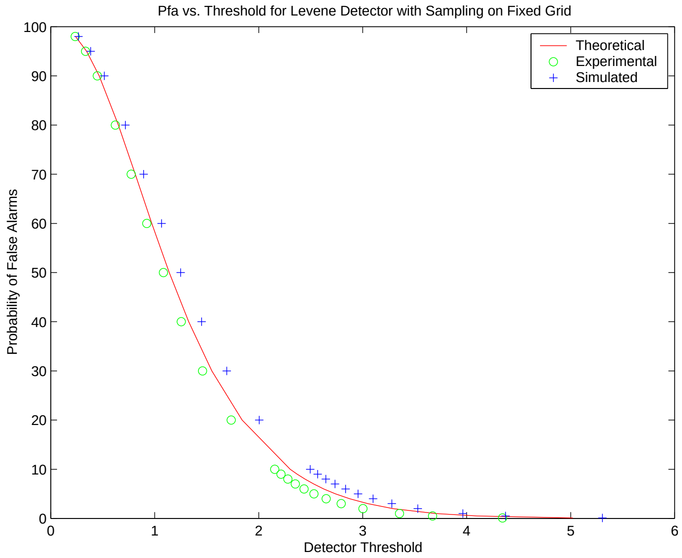


Figure 5.34: P_{fa} vs. threshold for Levene test for sampling on a fixed grid

Results

The detector based on the Levene test for equality of variances was tested on the image of fig 5.4. In fig 5.35 to fig 5.38 results of the detector are shown for 10% random sampling and applying a threshold of 4, corresponding to the empirical 1 % error rate level. Results are shown for different orientations of the scanning window.

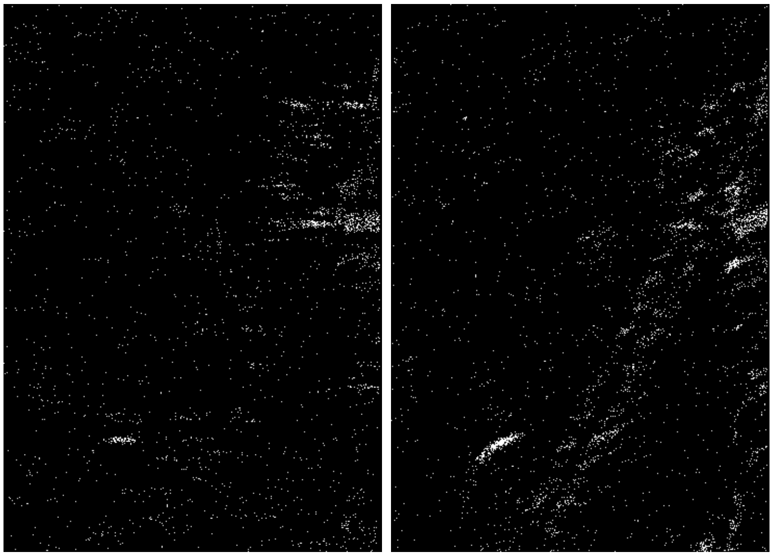


Figure 5.35: Results for orientation 0^o and 22.5^o

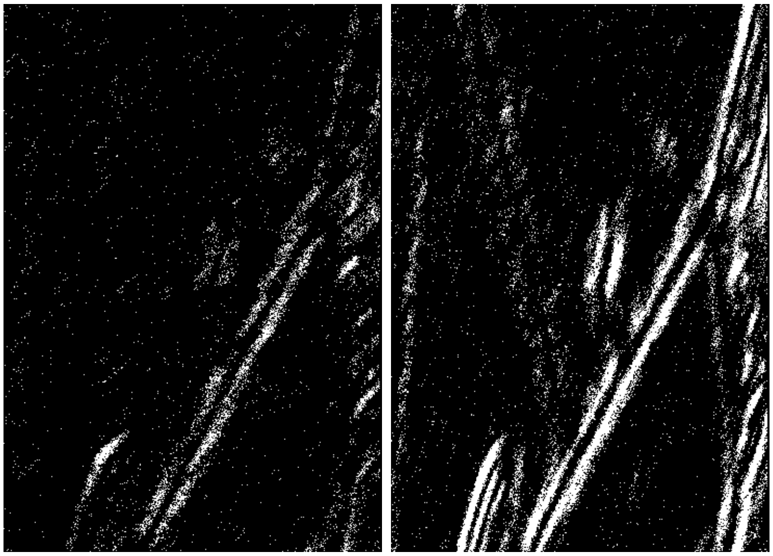


Figure 5.36: Results for orientation 45^o and 67.5^o

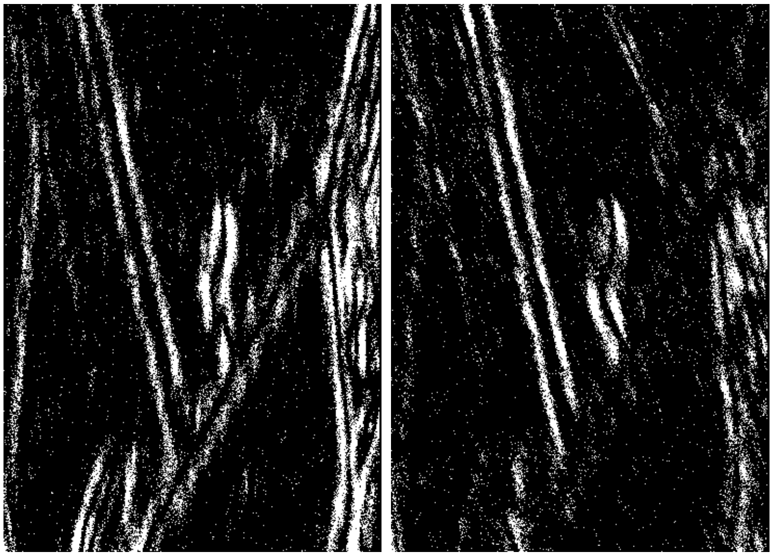


Figure 5.37: Results for orientation 90^o and 112.5^o

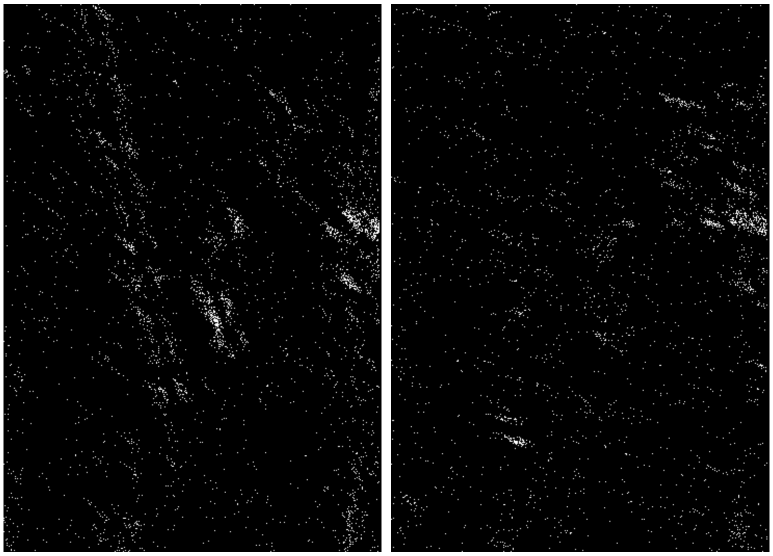


Figure 5.38: Results for orientation 135^o and 157.5^o

Fig. 5.39 presents the ROC curves for the Levene test obtained in simulated images with correlated speckle and containing a vertical edge of given contrast. Please note that the results are better than the results obtained for the Student test based detector (in fig. 5.18). This is probably due to the fact that the Levene test uses the three polarisations at the same time. Although the simulated images were constructed such that the three polarisation have the same variance in the SLC, the three polarisations do represent three different realisations of a normal distribution. Multi-variate detectors take advantage directly of this combined set of (almost) independent information.

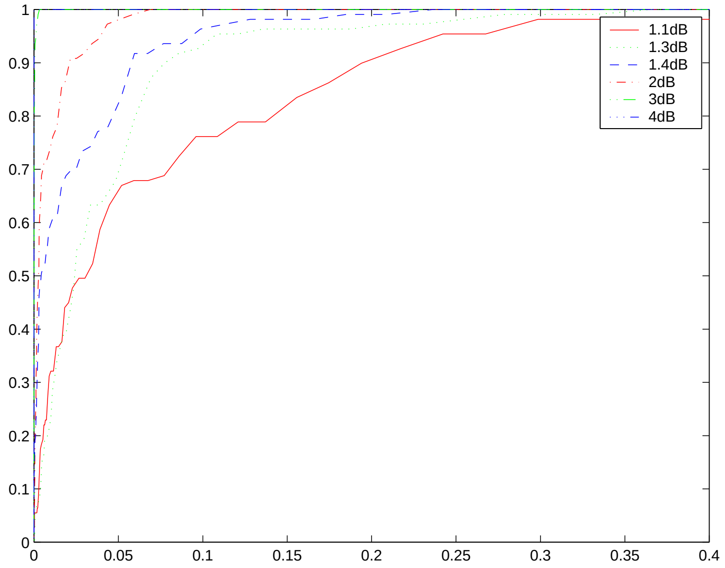


Figure 5.39: ROC curves for the Levene test at different contrast levels

5.4.2 Edge Detection based on a Difference in Means

Principle of the Detector and Test Statistics

The idea is to apply a multi-variate test for difference in means between two populations. The test will be applied to the log-intensity images where differences in radar reflectivity appear as differences in central tendencies such as the mean. The Hotellings T^2 test [36, 39] that was mentioned in the previous section was used.

The *Hotellings T^2-statistic* is defined as:

$$T^2 = \frac{n_1 n_2 (\overline{\mathbf{X_1}} - \overline{\mathbf{X_2}})^t \mathbf{C}^{-1} (\overline{\mathbf{X_1}} - \overline{\mathbf{X_2}})}{n1 + n2}, \tag{5.25}$$

with $\overline{\mathbf{X_i}}$ the vector of averages of the values in scanning window i. Each element of the vector represents the average for one polarisation. $[C]$ is the pooled covariance matrix as defined in eq. 5.20.

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

$$F = \frac{(n_1 + n_2 - p - 1)T^2}{(n_1 + n_2 - 2)p} \tag{5.26}$$

follows an F distribution with degrees of freedom p and $(n_1 + n_2 - p - 1)$. This time, the number of variants, $p = 3$ because we are dealing with log-intensity images and we only have 1 image for each polarisation.

The hypotheses that need to be full-filled in order that this test should be valid are:

- the distribution in both populations to be compared should belong to the same family

- the sample size from both regions should be similar

- the sample size should be sufficiently large (> 30)

In uniform (non-textured) regions the first condition is full-filled. The second condition is full-filled by construction of the detector, i.e. we pick the same number of pixels from both rectangular windows. The third condition limits the size of the windows or the sampling rate within the windows. The log-intensity image is the best candidate for comparing regions on the basis of this test because differences in radar reflectivity appear in these image purely as variations of means and because the form of the distribution is independent of the radar reflectivity.

The theoretical F-distribution is now $F(3, n_1 + n_2 - 3 - 1)$ which yields theoretical 1% and 5% false alarm thresholds of 3.97 and 2.69 respectively for a sampling rate of 10 % in a 51×11 window.

Pct	N	ν_2	5 % Threshold of $F(3, \nu_2)$	1 % Threshold of $F(3, \nu_2)$
5	28	52	2.783	4.182
10	57	110	2.687	3.965
20	114	224	2.645	3.870
30	171	338	2.631	3.840
40	228	452	2.625	3.825
50	286	568	2.621	3.816
60	343	682	2.618	3.810
70	400	796	2.616	3.806
80	457	910	2.615	3.803
90	514	1024	2.614	3.801
100	572	1140	2.613	3.799

Table 5.5: Theoretical 5% and 1% false alarm threshold for F-distribution vs. sampling percentage in a 51x11 window

Application on simulated images without correlation

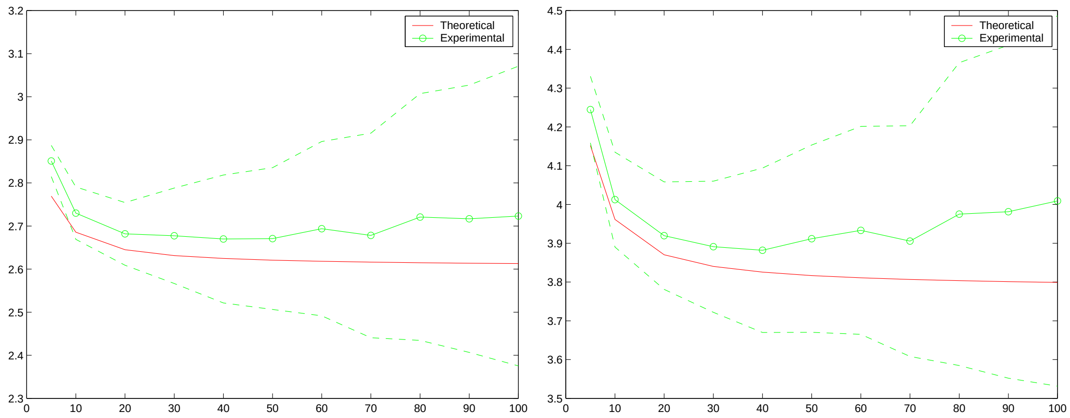


Figure 5.40: 5% (left) and 1% (right) threshold vs. sampling ratio for sampling without replacement (in simulated images without spatial correlation)

In simulated images without spatial correlation the threshold corresponds to the theoretically predicted one. This is illustrated in fig. 5.40.

Application on uniform parts of the SAR image (with correlation)

In fig. 5.41 the 5 and 1 % thresholds are plotted versus the sampling ratio. The figure shows the averaged results for 15 uniform regions of the log-intensity image. The theoretical threshold is lower than the one that was found experimentally for all sampling ratios. In fig. 5.42 the 5 and 1% thresholds are shown as a function of the sampling ratio for the 9 uniform regions with a spatial correlation compatible with the SAR system's ACF. The difference in threshold between vertical and horizontal windows is only negligible at 10 % sampling ratio. For the 9 regions compatible with the average ACF, the thresholds found at 100 % do correspond to the corrected thresholds presented in table C.2 of the annex and repeated in table 5.6.

5 % FA Threshold		1 % FA Threshold	
Hor.	Vert.	Hor.	Vert.
12.49	13.14	18.15	19.10

Table 5.6: Theoretical 5% and 1% false alarm thresholds taking into account spatial correlation for the Hotellings test

For these regions, thus for non-textured regions the relative variation of the threshold over the different regions is independent of the sub-sampling ratio.

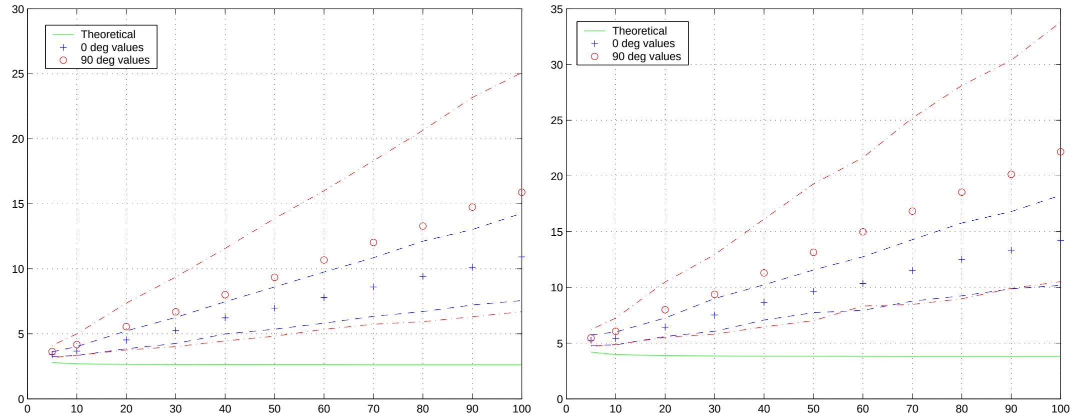


Figure 5.41: 5% (left) and 1% (right) threshold vs. sampling ratio for horizontal and vertical windows for the 15 homogeneous regions

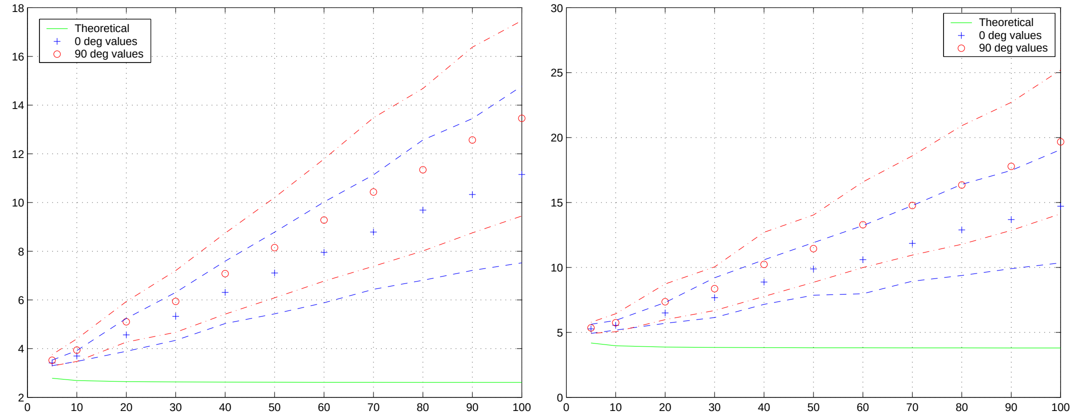


Figure 5.42: 5% (left) and 1% (right) threshold vs. sampling ratio for horizontal and vertical windows for the regions compatible with the ACF

It is thus again not possible to use the theoretical F-distribution to find the CFAR threshold. In order to find the thresholds we have determined the threshold as a function of the probability of false targets for 15 image windows with varying brightness (and for the 8 orientations of the rectangles). In fig. 5.43 the probability of false alarms is plotted against the average threshold. The curves of average minus and plus one standard deviation are also shown as well as the theoretical curve. For the 15 regions we simulated also images with the same radar reflectivity and the average spatial correlation found in uniform regions. The difference between the simulated curve and the experimental one is due to the fact that some of the selected regions in the real image do contain some texture, increasing their spatial correlation. The curves were determined based on a 10 % sampling ratio.

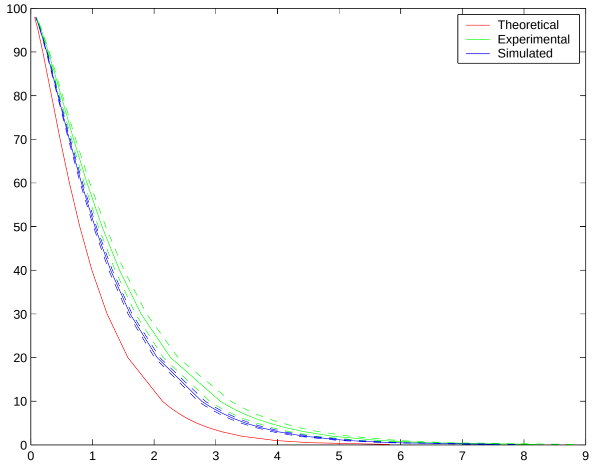


Figure 5.43: Probability of false alarms vs. detector threshold for the Hotellings-T^2 test

The variation due to radar reflectivity and scanning window orientation is very small. However, the probability of false targets is always higher than the theoretically predicted one. The CFAR thresholds are determined experimentally from the simulated windows with correlation as for the Levene Test.

Sampling on a Fixed Grid

For sampling on a fixed grid the experimentally found thresholds in uniform regions of the SAR image and in simulated images with correlated speckle are slightly above the theoretical thresholds multiplied with the appropriate correction factor for the single-look log-intensity image presented in table C.3 of the annex. The correction factor used in the figure was 1.048 which is the average value for vertical and horizontal windows.

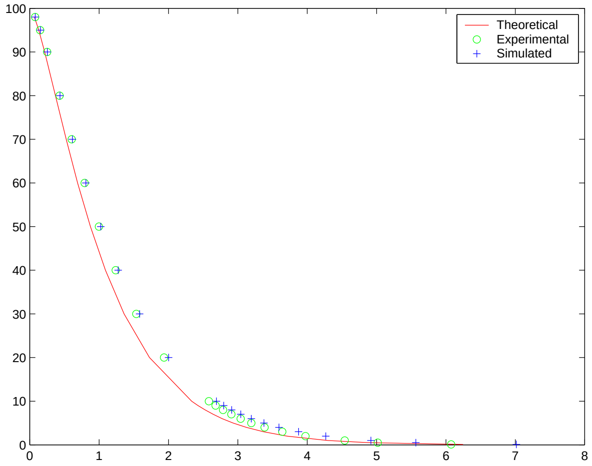


Figure 5.44: P_{fa} vs. threshold for the Hotellings test for sampling on a fixed grid

Results

Figure 5.45 shows the ROC curve for the Hotellings test at various contrast-levels.

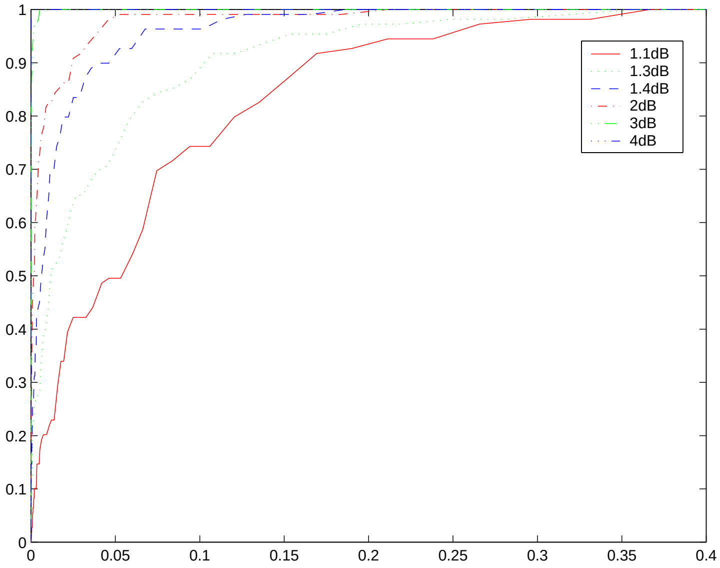


Figure 5.45: ROC curves for the Hotellings test at different contrast levels

5.5 Fusing the Results for Different Orientations

Up-to now we have been working with the responses of the edge detectors for each orientation of the scanning window separately. In order to obtain a final edge map the results for the different orientations need to be combined. Several ways to do this are suggested in [23]:

- Sum: Summing the results from all orientations which is equivalent to the taking average.
- Maximum: Keep the maximum over all orientations
- Generalised Norm: Expand the norm to multiple orientations:

$$E_f = \sqrt{\frac{1}{4}(E_{0^o}^2 + E_{90^o}^2) + \frac{1}{4}(E_{45^o}^2 + E_{135^o}^2) + \frac{1}{4}(E_{22.5^o}^2 + E_{112.5^o}^2) + \frac{1}{4}(E_{67.5^o}^2 + E_{157.5^o}^2)}, \tag{5.27}$$

  with E_{α^o} the response of the detector for orientation α of the scanning windows.

A ROC curve was constructed for these 3 combination methods for the Levene and the Hotellings test. The ROC curves were calculated based on the results obtained in the real test image (fig. 5.4). In fig. 5.46 the results are shown. For the performance measured by these ROC curves the maximum gives the best results. We will thus use the maximum to combine the results from the different directions.

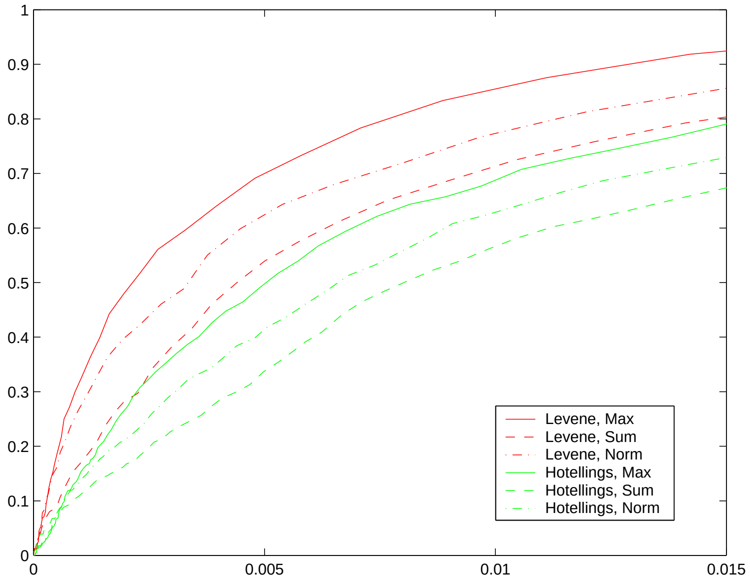


Figure 5.46: ROC curves for the different methods to combine the directional responses

The results of applying the 3 methods to the outcome of the Levene detector are shown in fig. 5.47.

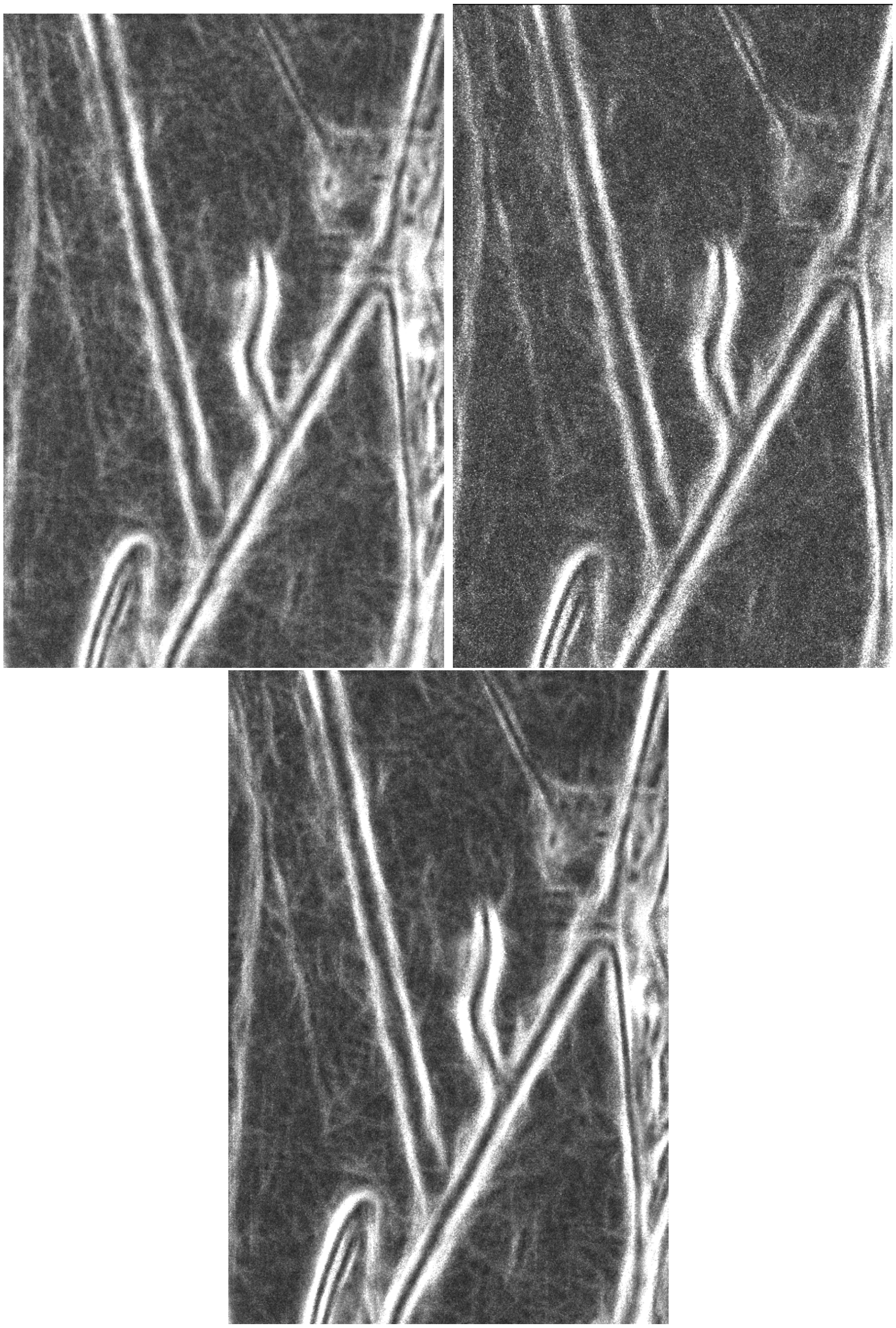


Figure 5.47: Results of combining the different Orientations of the Scanning Windows. Upper Left: Sum, Upper Right: Maximum, Bottom: Generalised Norm

5.6 Post-Processing: Improvement of Edge Localisation

As mentioned in the introduction to this chapter, the width of the response of the different detectors depends on the width of the used windows. This results in blurred, wide edges (see fig. 5.47). In order to sharpen these edges some filters were applied. As the individual detectors are 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 filters, i.e. one dimensional filters are applied in a direction perpendicular to the edge. These filters have therefore to be applied before combining the results for the different orientations of the scanning windows. We have tested three types of filters to improve edge localisation. They are described in the next sections.

5.6.1 Bartlett Filter

A Bartlett filter can be applied directly on the output from the detectors for a given edge direction. The Bartlett filter $f_B[k]$ is a triangular filter defined as:

$$f_B[k] = \begin{cases} \frac{2(k-1)}{n-1} & 1 \leq k \leq \frac{n+1}{2} \\ \frac{2-2(k-1)}{n-1} & \frac{n+1}{2} \leq k \leq n \end{cases}, \tag{5.28}$$

where n is the dimension of the filter. The above definition is for n odd. For n even the upper limit for the first case and the lower for the second are respectively $n/2$ and $n/2+1$. As we wish to have a symmetric filter around the current point, n will always be odd. The Bartlett filter results in a narrowing of the edges if n is chosen appropriately.

Results of applying this filter on the output of the Levene detector and taking the maximum of all orientations of the scanning windows are shown in figure 5.48.

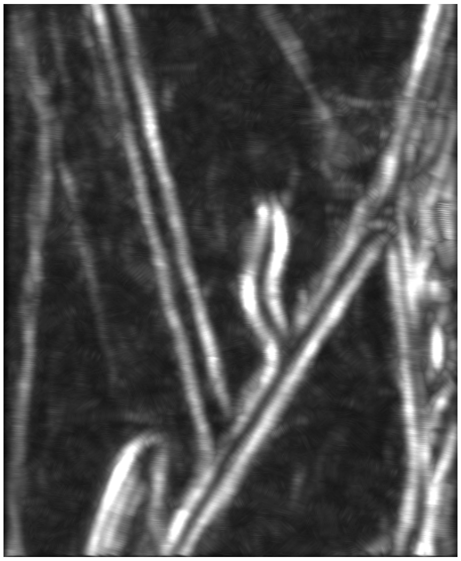


Figure 5.48: Results after applying Bartlett filter

5.6.2 Binary Skeleton

If we know the 5 % (or 1 %) false alarm threshold, that threshold can be applied in order to obtain a binary image. The binary skeleton filter finds the centre of binary blocks of given size and replaces its value with the maximum filter response inside that block, setting the rest to zero. 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. A 3×3 kernel was used for the morphological operator.

Results of applying this filter on the output of the Levene detector and taking the maximum of all orientations are shown in figure 5.49 with (right) and without (left) the use of the morphological operator.

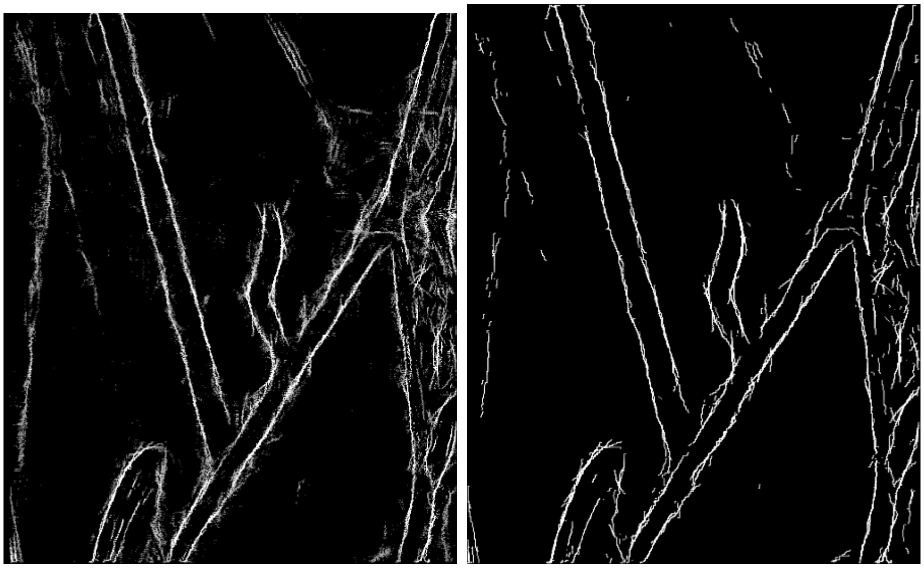


Figure 5.49: Results after applying the Binary Skeleton with (right) and without (left) using the morphological operator

5.6.3 Morphological Filter

Another way to post-process the results of the detectors is based on morphological operations with directive kernels. In fig. 5.50 a flow chart of the method is shown. The method is partly based on an article by Chanussot [41].

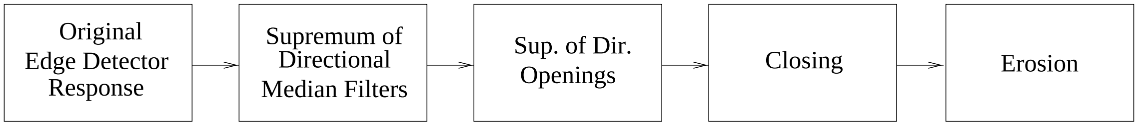


Figure 5.50: Flow chart of the morphological filter

The aim of the first step is to fill small gaps in the contour detector's result image. This could be achieved by a closing with reconstruction as is proposed by Chanussot [41]. However, we have obtained better results using a median filter [42]. The idea is to calculate the median in each pixel inside an elongated rectangular array around the current pixel. The maximum response will be obtained if the rectangle is oriented in the direction of the edge. The maximum is determined over a small range of orientations around the orientation of the detector's masks. The current pixel is then replaced by this maximum value of the median. The first step of the morphological is thus the "supremum of directional medians". This step combines smoothening with the filling of small gaps.

The next two steps do follow the approach presented in [41]. The supremum of directional openings is used to remove nonlinear or too short mountains in the detector's response. A simple closing with a small (5×5) structuring element is used to remove isolated false alarms.

The last step in [41] is a top-hat to remove linear mountains that are too wide. We replaced this by an erosion to further narrow the result. In fig. 5.51 the results of the successive steps of this morphological filter are shown on a small (32×32) part of the test image. The selected image part contains one true edge and is represented as a 3D surface in the figure.

In fig. 5.52 the result of applying the morphological filter to the outcome of the Levene detector is shown after taking the maximum of the result of the 8 orientations of the scanning windows.

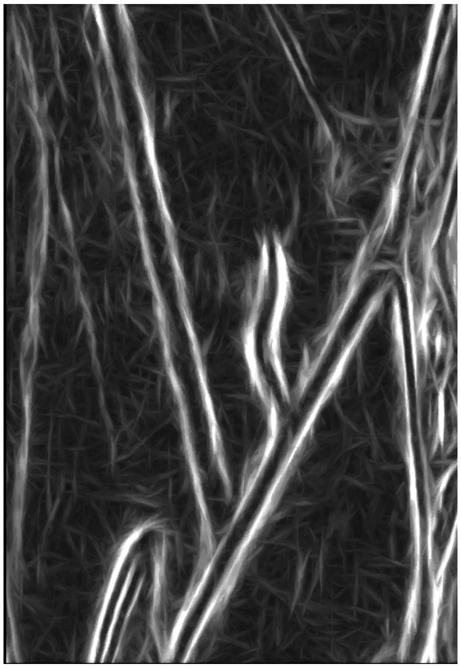
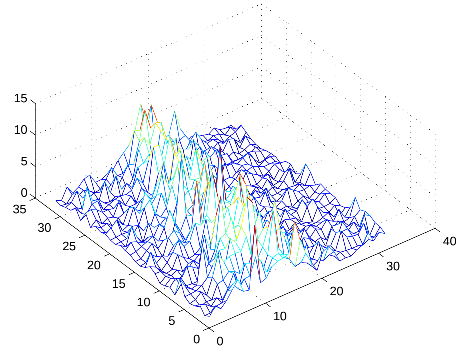
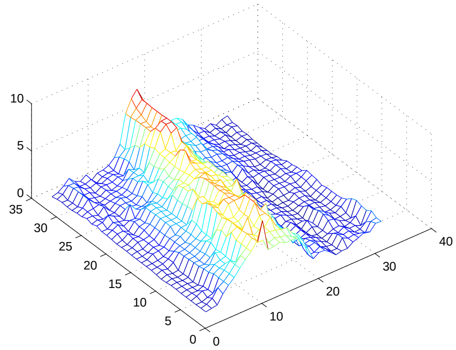


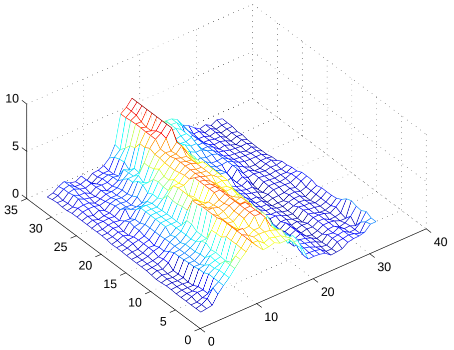
Figure 5.52: Results after applying the "Morphological Filter"



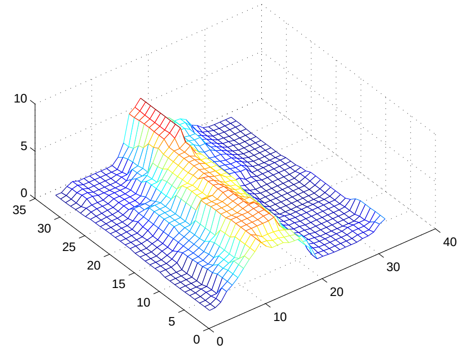
Orig. Edge Response



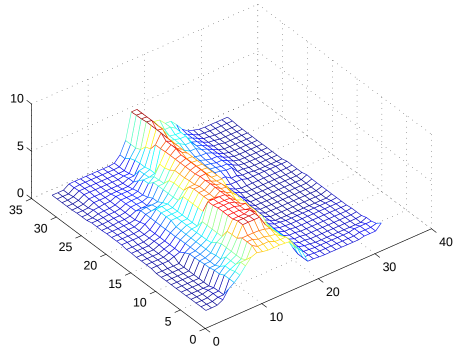
Sup. Dir. Medians



Sup. Dir. Openings



Closing



Erosion

Figure 5.51: Results of successive steps of the morphological filter

5.6.4 Vectorisation of the results

An existing method for extracting bright bars (lines) and converting them into objects [31] is used to vectorise the results of the edge detectors. The bar detector is applied to the thresholded results of the detector either with or without the improvement of edge localisation. Results of the vectorisation applied to the results of the binary skeleton filter are shown in fig 5.53.

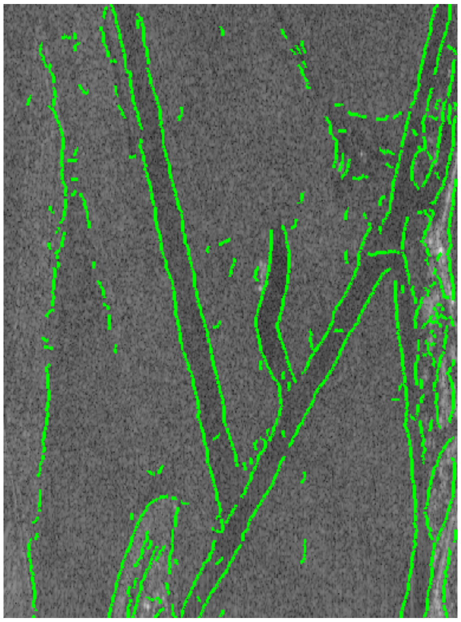


Figure 5.53: Vectorised edges

5.6.5 Conclusions about Edge Localisation

The first two methods for improving the localisation of the edges just find the centre of the detector's response. The binary filter is more effective than the Bartlett filter. The third method, based on morphological operators, locates a ridge of maximum response of the detector. These three methods are relatively fast but they do not necessarily give the correct position of the edge.

All of the presented method are based on the results of a SWCE (Scanning Window Central Edge) edge detector. In fact edge detectors based on the SWCE principle result in a localisation bias towards the side of the darkest region [43, 44]. For our purposes the results obtained above are sufficient. If a more precise edge localisation is needed, methods that combine the results of an edge detector with local measures in the original image and a regularisation method that avoids to follow local speckle patterns should be used. Germain [44, 45, 46] proposes a method based on active contours. Fjortoft [23, 35] combines an edge detector with a watershed method as well as a method based on Markov Random Fields. All these methods are very computationally intensive but they allow to fill gaps in the edge map in addition to improving edge localisation. Oliver [34] uses a Fixed Window Scanning Edge (FWSE) approach to localise the position of edges more precisely. His method is in fact still biased and can not cope with edges that are not exactly aligned with the orientation of the scanning windows.

5.7 Development of a Fusion Strategy for Edge Detectors

All detectors were run on the test image of fig. 5.4. In order to combine the results for the eight orientations of the scanning windows the maximum over all orientations was determined. To avoid biasing the results by the quality of the different thresholds or by any further post-processing, the ROC curves were determined directly on this new edge image. Random subsampling at 10 % sampling ratio was used for all detectors. The results are presented in fig. 5.54. The results of the Touzi detector [2] are also shown. Note that in these ROC curves we have set $\Delta_{max} = 2$ (see section 5.1 for meaning of Δ_{max}).

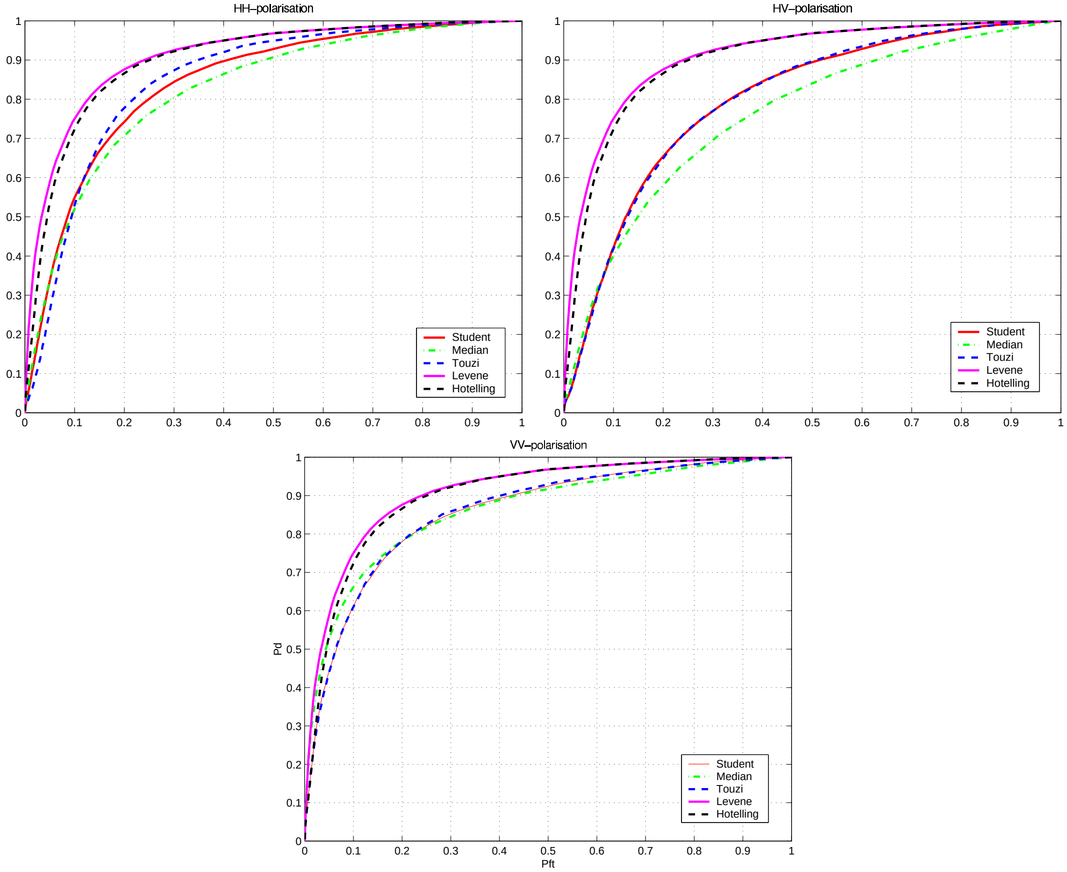


Figure 5.54: Comparison of the different contour detectors for the test image Left: HH-polarisation, Right: HV Polarisation, Bottom: VV-Polarisation

In fig 5.54 the results of the uni-variate detectors are presented for each polarisation in a separate graph. The results of the multi-variate detectors are of course those obtained using all three polarimetric images. Several conclusions can be drawn from fig. 5.54:

- the multi-variate detectors always outperform the uni-variate detectors when no fusion is applied. This means that taking into account the complete polarimetric information gives better results than only using the "best" polarisation separately.
- the best uni-variate detector seems to depend on the polarisation, i.e. on the local information content of the image.
- the best results of the uni-variate detectors are obtained in the VV-polarised image
- the Student test based detector gives results that are very close to the results of the Touzi detector.
- the Student test and the Touzi detector give better results than the median test except in HV where, at low false alarm rates, the median test has a slightly higher probability of detection

The reason why the multi-variate detectors always outperform the uni-variate detectors is that the different polarisation images contain complementary information. The multi-variate detectors take this information directly into account. In order to use the same information in the uni-variate detectors, their results need to be fused. This is examined in the next section. The two multi-variate detectors have been noted to give complementary results. Fusion of their results is therefore likely to improve the results of the edge detection. The fusion of the results of the two multi-variate detectors is examined in sect. 5.7.2

Fusion of the results of edge detection can be done at several levels:

- Fuse the resulting vectorised objects: this will involve a combination of voting mechanisms and higher level constraints such as continuity of edges
- Fuse just before the vectorisation on the thresholded image of the detector, i.e. after combining the results for the different orientations of the scanning rectangles
- Fuse the results for each orientation of the scanning rectangles separately: this automatically ensures that only edges, that were found to have a similar orientation, are fused. This is the method that we used.

5.7.1 Combining the Results for the Uni-Variate Edge Detectors

The uni-variate edge detectors are applied to the three polarimetric images separately. If an edge is present in the scene it is likely to find it in the three images. This yields 3 detections to be combined into a final decision. It is likely that, when several bands of data are available, a multi-variate method applied simultaneously to all bands gives better results than the combination of results from uni-variate detectors.

Fig. 5.55 presents the ROC curves of the Student-test based detector for random subsampling and sampling on a fixed grid. The results obtained for the three polarisations are shown.

Fig. 5.56 presents analogous results for the median test. In fig. 5.57 the results for the Touzi detector [2] (applied on the intensity images) are shown. For the test image the ROC curves obtained for VV and HH cross each other for each of the detectors. This means that HH and VV provide complementary information. The ROC curve for HV is below the one obtained for the other two polarisations.

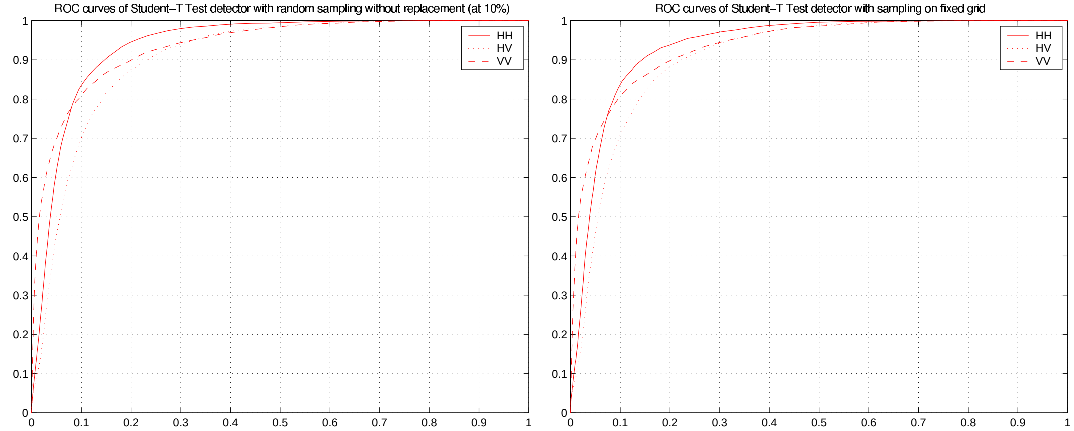


Figure 5.55: Influence of polarisation on the results of the Student-test detector (random subsampling (left) and sampling on a fixed grid (right)

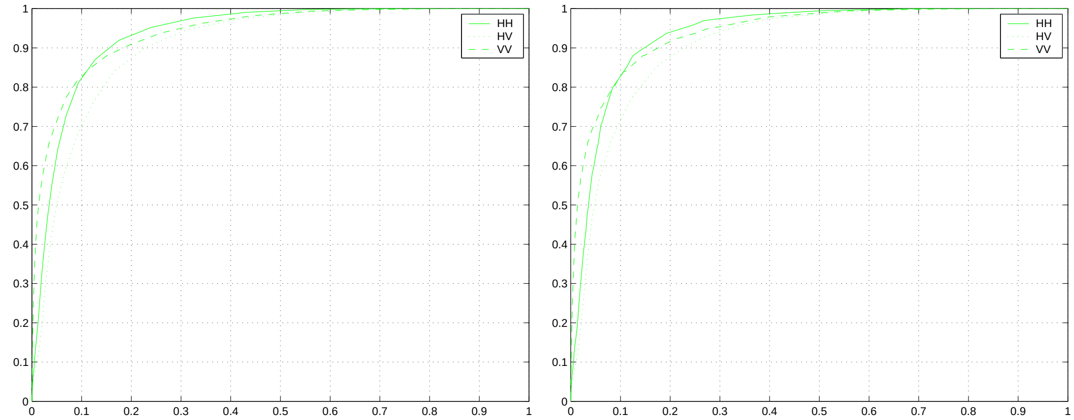


Figure 5.56: Influence of polarisation on the results of the median-test detector (random subsampling (left) and sampling on a fixed grid (right)

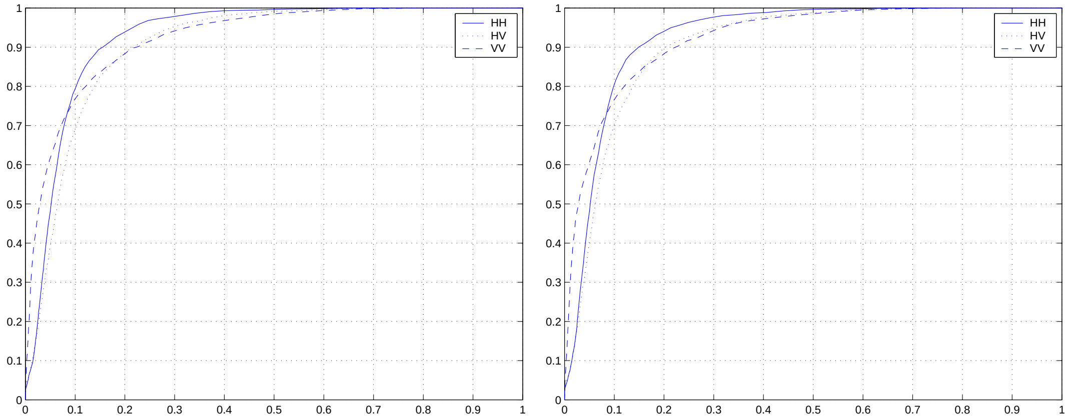


Figure 5.57: Influence of polarisation on the results of the Touzi detector (random subsampling (left) and sampling on a fixed grid (right)

Fusion by Arithmetical Operators

Arithmetic operators can be applied directly on the detector's outcome. An overview of information fusion operators is presented in a paper by I. Bloch [47]. Some of the operators discussed in her paper were investigated here.

Let X_{HH}, X_{HV} and X_{VV} be the response of a detector obtained in the three polarisations. The different operators that were used are then defined as:

- **Maximum:**

$$R_{Max} = MAX(X_{HH}, X_{HV}, X_{VV}) \tag{5.29}$$

- **Arithmetic Mean:**

$$R_{Mean} = \frac{X_{HH} + X_{HV} + X_{VV}}{3} \tag{5.30}$$

- **Symmetrical Sums:**

  The general form of a symmetrical sum is given by [47]:

$$\sigma(X_{HH}, X_{HV}, X_{VV}) = \frac{g(X_{HH}, X_{HV}, X_{VV})}{g(X_{HH}, X_{HV}, X_{VV}) - g(1 - X_{HH}, 1 - X_{HV}, 1 - X_{VV})} \tag{5.31}$$

  We experimented with two symmetrical sums:

- **Associative Symmetrical Sum:** This is obtained for
$$g(X_{HH}, X_{HV}, X_{VV}) = X_{HH}X_{HV}X_{VV} \tag{5.32}$$

- **Non-Associative Symmetrical Sum:** For which
$$g(X_{HH}, X_{HV}, X_{VV}) = X_{HH} + X_{HV} + X_{VV} - X_{HH}X_{HV}X_{VV} \tag{5.33}$$

For the symmetrical sums the values of X_{HH}, X_{HV} and X_{VV} need to be confined to the interval [0..1]. As the best results for the different edge detectors are obtained by imposing a threshold corresponding to a range between 10 and 0.1 % false alarms, this is the part of the curve that is used for the conversion to the interval [0..1]. The output values R of each detector are converted into a notion of "confidence in the presence of an edge" C using the relation between the detector's output and the probability of false alarms that would be obtained if that value was used as a threshold for the detector. The conversion to confidence that is used is defined as:

$$\begin{array}{ll} if(P_{fa}(R) \geq 0.1) then & C = 0, \\ else\ if(P_{fa}(R) \leq 0.001) then & C = 1, \\ else & C = 1 - P_{fa}(R). \end{array} \tag{5.34}$$

Fusion by Geodesic Reconstruction

In [41] J. Chanussot proposes a context dependent method, called geodesic reconstruction. The method is based on simple morphological operations.

The idea is that the ideal result of fusion lies between the maximum and the minimum response of the detector for the different channels (polarisations). If, at a given position, an edge is detected in all polarisations, the minimum and maximum are both high and close together. This gives a high confidence in the presence of an edge in that point. If such a point is found in the image it is likely that some neighbouring points also belong to an edge. Even if only one polarisation would confirm this. The proposed fusion thus starts from the minima as markers that are dilated until they coincide with the maximum, i.e. starting from the minima of all detector's responses a dilation is iteratively executed until idempotence with the maximum. The method is illustrated in fig. 5.58.

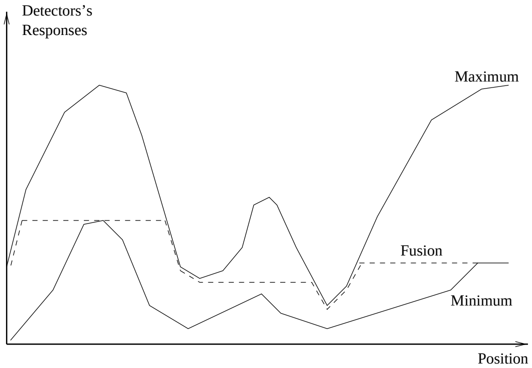


Figure 5.58: Principle of fusion by geodesic reconstruction

Fusion based on Fuzzy Logic

An alternative fusion approach is based on fuzzy logic . We would like to associate a low confidence in the presence of an edge to a low value of the outcome of the edge detector in each pixel of the image. A low value of the outcome means that, if we reject the null-hypothesis for that value (i.e. we declare that the corresponding pixel belongs to an edge), we have a high probability of making an error, thus a high probability of false alarms.

The confidence in the presence of an edge can thus be linked to the probability of false alarms. The input membership function for our fuzzy-logic-based approach are based on this fact. In fig. 5.59 the general input membership function that was used is shown. T_1, T_2, T_3 are the thresholds that correspond to a given probability of false alarms for the detector. These thresholds are derived from the relation between the outcome of each detector and the probability of false alarms for each of the detectors (given in figs. 5.11 and 5.23).

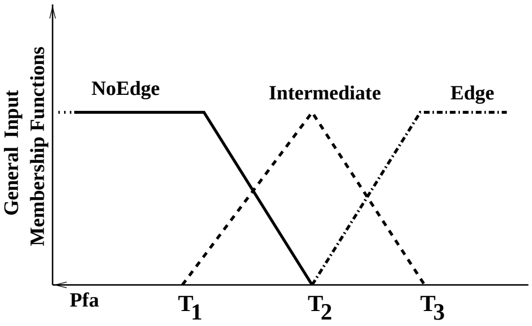


Figure 5.59: General input membership functions for edge detectors

To determine the actual input membership functions for each detector we have used the relation between the detector's threshold and the probability of false alarms. The values used for determining the boundaries of the membership functions are given in table 5.7. The best results for the fuzzy-logic-based fusion were obtained by setting $T_1 = T(P_{FA} = 10\%), T_2 = T(P_{FA} = 1\%), T_3 = T(P_{FA} = 0.1\ \%)$.

Detector Name	Probability of False Alarms				
	10 %	5 %	1 %	0.5 %	0.1 %
Difference of Means	1.87	2.24	2.95	3.22	3.78
Difference of Medians	2.83	4.05	6.97	8.35	11.55
Touzi Detector	0.30	0.35	0.43	0.46	0.51

Table 5.7: CFAR thresholds for the uni-variate detectors

For the output of the fuzzy inference system, five output classes were defined. They are shown in fig. 5.60

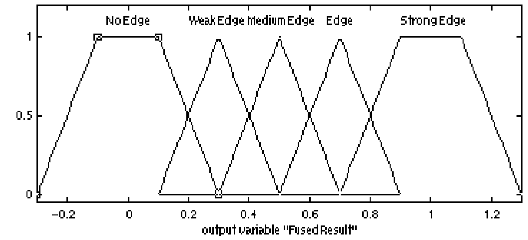


Figure 5.60: Output membership functions

The used rules are summarised in table 5.8. The combination and implication operators were both MIN. The centroid defuzzification method was used.

Polarisation 1	Polarisation 2	Polarisation 3	Fused Result
NoEdge	NoEdge	NoEdge	NoEdge
NoEdge	NoEdge	Intermediate	WeakEdge
NoEdge	NoEdge	Edge	MediumEdge
NoEdge	Intermediate	Intermediate	WeakEdge
NoEdge	Intermediate	Edge	MediumEdge
NoEdge	Edge	Edge	MediumEdge
Intermediate	Intermediate	Intermediate	MediumEdge
Intermediate	Intermediate	Edge	Edge
Intermediate	Edge	Edge	StrongEdge
Edge	Edge	Edge	StrongEdge

Table 5.8: Fuzzy rules for combining the results of the uni-variate edge detectors for the three polarisations

Results of the Fusion of the Uni-Variate Edge Detectors

In fig. 5.61 the ROC curves are shown for the different fusion methods applied to the results of the Student-test based detector (left) and the median test (right). The curves for the average is completely equivalent with the non-associative sum. The maximum operator gives slightly better results. The fuzzy-logic based approach and the geodesic reconstruction gives the poorest results. The best results obtained by fusing the response of the uni-variate Student-test based detector over the three polarisation are still less good than than results of the multi-variate detectors which are also shown in the figure. For the Median test the order of the results of the different fusion methods is similar to that obtained for the Student test but the fused results are equivalent to those obtained by the Hotellings test.

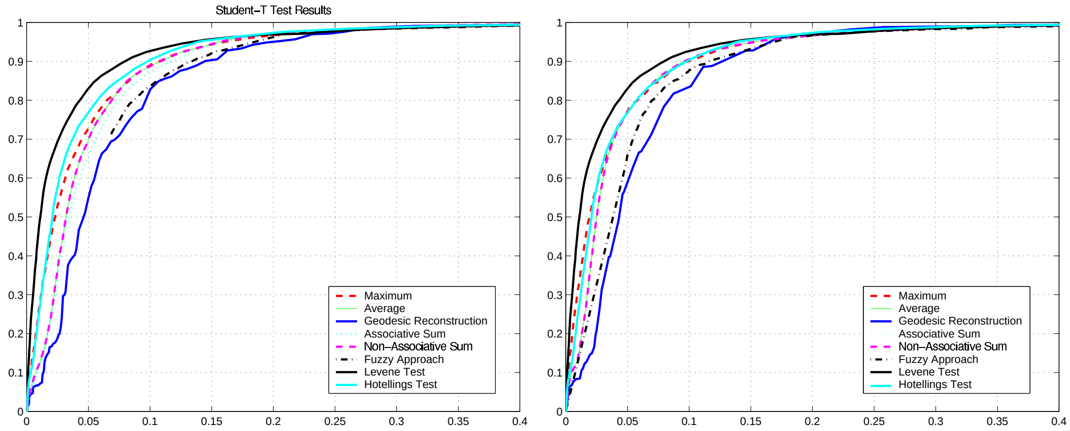


Figure 5.61: ROC curves for the different fusion methods of the results of the Student-test based (left) and the median-test based detector (right). Both for sampling on a fixed grid

5.7.2 Fusion Strategy for Multi-Variate Edge Detectors

In the previous section we have seen that fusing the results of the uni-variate detectors results in a performance that is in the best case equal to the performance of the multi-variate detectors. In this section we will investigate whether it is possible to improve the results of edge detection by combining the results of the two multi-variate detectors (i.e. the Levene and the Hotellings test). The results of the edge detectors based on the Levene and Hotellings test are given in figure 5.62 and 5.63. They appear to be complementary. The Hotellings test is able to detect smaller roads while some of the edges of large objects are delimited better by the Levene test.

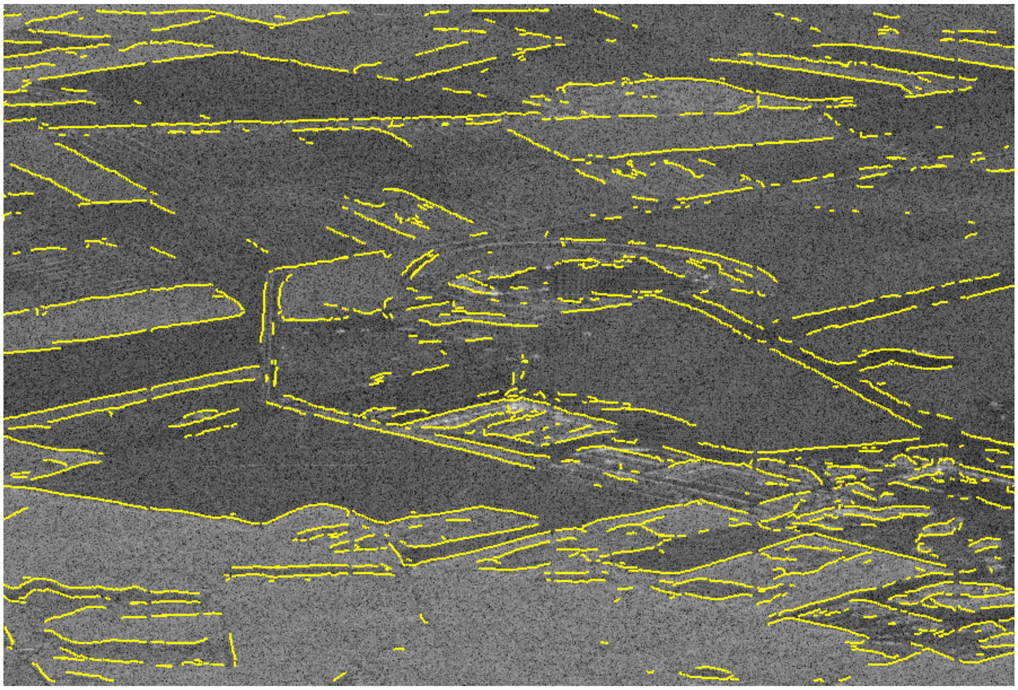


Figure 5.62: Results of Levene detector

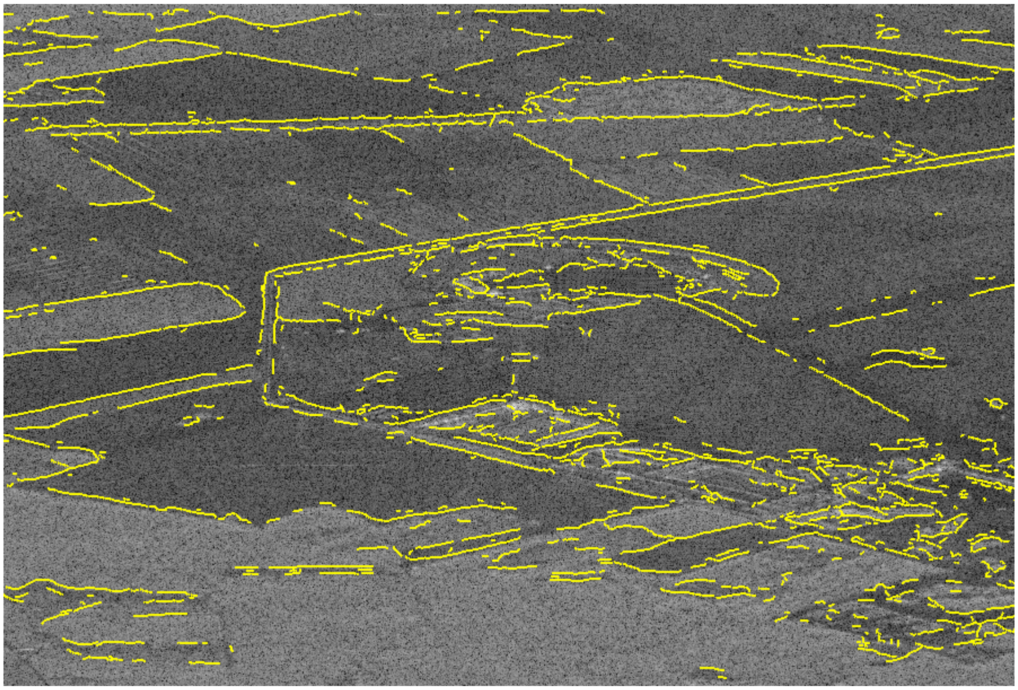


Figure 5.63: Results of Hotellings detector

Another indication of the complementarity of both detectors is given by the ROC curves. In figure 5.64 the ROC curves for both detectors are shown for simulated images with a vertical edge at various contrast levels. From the ROC curves it can be concluded that neither of the two detectors is best over the complete range, i.e. the curve for the Levene test crosses the one for the Hotellings test and the point where this crossing occurs depends on the contrast. This means that, even if the false alarm rate is fixed, it is not possible to say that one detector is better than the other one, because this will depend on the contrast of the edges in the image. Because both detectors use the complete inter-channel covariance matrix, the inter-channel correlation within a region also influences the behaviour of the detectors. This is illustrated by fig. 5.65. To obtain the ROC curves shown there, two images were simulated with a vertical edge having the same contrast. However in the first image the inter-channel correlation was set to zero (for both halves of the image) and in the other the correlation between HH/HV and HH/VV were both set to 0.9. At high inter-channel correlation, the Hotellings test gives the best results while at the low inter-channel correlation the Levene test seems to give slightly better results. It is thus not possible to decide a priori which test is better and it is therefore useful to fuse the results of both detectors.

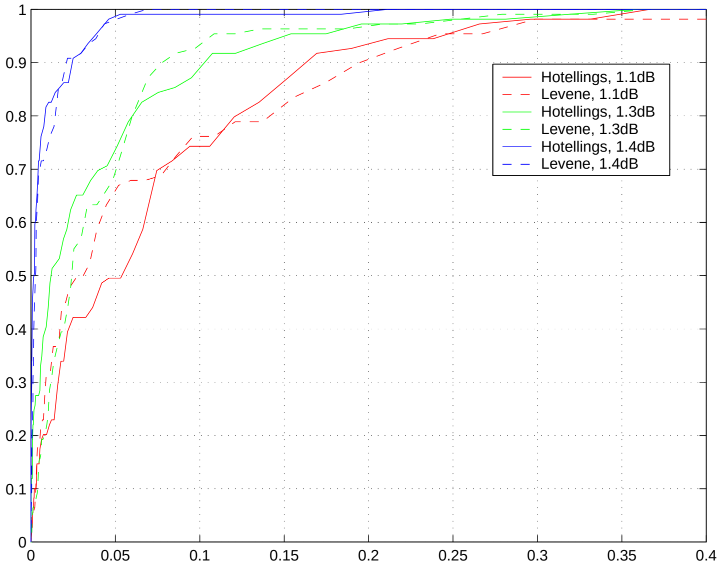


Figure 5.64: ROC curves for Levene and Hotellings test at different contrast levels

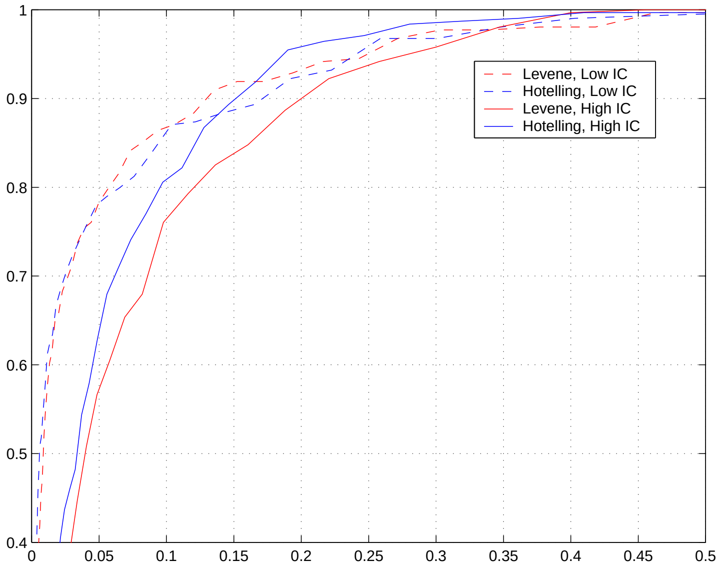


Figure 5.65: ROC curves for both tests at constant contrast level but varying inter-channel correlation

The benefit obtained by the fusion can lie in the increase of the probability of detection, the decrease of the probability of false alarms or both. The effectiveness of the fusion not only depends on the fusion algorithm, but also on the range of outcomes of the individual detector's responses. We have already seen that the two multi-variate detectors are complementary. In the next section we will investigate this complementarity further. In particular we will try to establish for which ranges of thresholds the results of both detectors can be enhanced significantly by fusion and what type of benefit can be expected from this fusion. Both edge detectors are considered as experts that give an opinion regarding the presence of an edge in each pixel of the image. The analysis below is based on the statistical comparison of these two experts.

Comparison of the 2 experts

The comparison of the two experts is based on their results on two small regions of the SAR image. The first is the region that was used as a test image in this chapter (see fig. 5.4), the second region is shown in fig. 5.66.

Figure 5.66: Second region of E-SAR used for the comparison of the two experts (HH: Red, HV: Green, VV: Blue)

On the two regions the ground truth was delimited manually, i.e. the edges were indicated. This ground truth is used to determine the probability of detection and false alarms for a given threshold for each detector. This information is used as input for the statistical methods that are used in the comparison. The statistical methods are described below.

The two detectors are considered as two experts yielding a binary decision (edge,no-edge) at every pixel of the test images. The decision is based on the choice of a pair of thresholds (T_{Levene}, $T_{Hotellings}$). This decision can be validated using the ground truth and from this validation two 2×2 matrices (contingency tables) are determined that form the basis for the statistical comparison of the two experts. One matrix describes the performance in terms of detection, the other one in terms of false alarms. Both matrices have the form given in table 5.7.2. The elements of the matrix are determined for all pixels not belonging to an edge (matrix M_{FA}) and for all points of the ground truth (matrix M_{Pd}), in a similar way as for the creation of the ROC curves mentioned before (see sect. 5.1).

If the two detectors agree completely, the elements n_{21} and n_{12} of the matrix are zero. If this is true for both matrices, M_{FA} and M_{Pd}, fusion will not be useful and any of the two experts can be chosen. If n_{21} and n_{12} are non-zero and almost equal, the two experts are providing complementary information and fusion is useful. The symmetry of the matrix is thus a first characteristic that gives an idea of the usefulness of fusion. The second char-

		Levene Test		
		No-Edge	Edge	
Hotellings Test	No-Edge	n_{11}	n_{12}	n_{10}
	Edge	n_{21}	n_{22}	n_{20}
		n_{01}	n_{02}	n

acteristic is the interdependence of the two experts. If the two experts are independent, the values n_{11} and n_{22} should be close to the values predicted from the marginal probabilities. More formally the two characteristics can be assessed using two statistical tests. The symmetry is described by the Mc Nemar test and the interdependence is investigated by the Kappa test.

The Mc Nemar Test

The Mc Nemar test [38] is a hypothesis test that was designed to test for the significance of changes, e.g. before and after a medical treatment. In our case it will thus test whether the opinions of the two experts differ significantly. The hypotheses are defined as:

$$\begin{array}{ll} H_0: & p_{21} = p_{12} \\ H_1: & p_{21} \neq p_{12} \end{array} \tag{5.35}$$

with e.g. $p_{12} = Prob(Hotellingsresult = No-Edge, Leveneresult = Edge)$. If the null-hypothesis is verified the expected value $e_{12} = e_{21} = \frac{n_{12}+n_{21}}{2}$. The test statistic:

$$\begin{aligned} d^2 &= \frac{(n_{12}-e_{12})^2}{e_{12}} + \frac{(n_{21}-e_{21})^2}{e_{21}} \\ &= \frac{(n_{12}-n_{21})^2}{n_{12}+n_{21}} \end{aligned} \tag{5.36}$$

follows a $\chi^2(\nu = 1)$ distribution.

Intermezzo: Example of the use of the Mc Nemar test

A candidate X for maior in a small village has spend a lot of money in his election campaign and wishes to know whether the money was well spend. A week before the campaign there was an election poll and people were asked for which of the two candidates (X or Y) they intended to vote. These results are combined with the results of a similar poll after the campaign to create a contingency table.

		Before campaign	
		X	Y
After Campaign	X	440	50
	Y	30	450

In this case $d^2 = 338.9$ which means that the campaign has had an influence on the behaviour of the voters.

The Kappa Test

This test, introduced by Cohen [48], is applied to the elements of concordance between the two experts, i.e. the elements n_{11} and n_{22} of the contingency table. It compares the values of these elements with the value the elements would have if the two experts were independent:

$$\begin{array}{ll} H_0: & p_{ii} = p_{0i}p_{i0} \\ H_1: & p_{ii} \neq p_{0i}p_{i0} \end{array} \tag{5.37}$$

The observed concordance is

$$p_0 = \frac{n_{11} + n_{22}}{n}, \tag{5.38}$$

the expected value is:

$$p_e = \frac{\frac{n_{10}n_{01}}{n} + \frac{n_{20}n_{02}}{n}}{n}, \tag{5.39}$$

The Kappa coefficient is defined as: $\kappa = \frac{p_0 - p_e}{1 - p_e}$. It can be shown that if n is large enough $(n > 8)$, the variable

$$Z_\kappa = \frac{\kappa}{\frac{\sqrt{p_e}}{\sqrt{n(1-p_e)}}} \sim \mathcal{N}(0,1) \tag{5.40}$$

For a given threshold $z_{\alpha/2}$ three cases are distinguished:

$$\begin{array}{ll} Z_\kappa < -z_{\alpha/2} & \text{two experts are discordant (they disagree)} \\ Z_\kappa > z_{\alpha/2} & \text{two experts are concordant (they agree)} \\ -z_{\alpha/2} < Z_\kappa < z_{\alpha/2} & \text{two experts are } \textit{non-concordant} \\ & \text{(any agreement or disagreement can be considered a coincidice)} \end{array} \tag{5.41}$$

Intermezzo: Example of the use of the Kappa test

Suppose physicians have been used to diagnose a given type of disease on an X-ray photograph and now a company has developed an environment for digitising the images and let the physicians play with the contrast interactively in order to reach their opinion. One would like to know whether using the new method influences the decision of the physicians. A contingency table is constructed from the decisions of a set of physicians using the old and the new method and the result is:

		Old method	
		No	Yes
New Method	No	48	2
	Yes	4	43

In this case $\kappa = 0.8956$ and $Z_\kappa = 8.7519$ and the two “experts” are thus concordant; the company will be able to sell its product.

Combining the two statistical tests

The two tests can be applied for a set of combinations of thresholds for the Levene and Hotellings test. For each pair of thresholds the contingency tables are determined, in test images, for false alarms and detections and the tests are applied. Combining the two tests gives, in principle, 6 possible values. They are listed below:

The discordant case was not observed in the test images ($Z_\kappa > 0$ for all thresholds). The non-symmetric non-concordant case corresponds to the case where the threshold for

McNemar Test	Kappa Test	Consequence for Fusion
Symmetric	Concordant	Fusion useful if $n_{12} > 0$ and $n_{21} > 0$
	Non-Concordant	Fusion not useful
	Discordant	NA
Non-Symmetric	Concordant	Fusion may be useful
	Non-Concordant	Choose the best expert
	Discordant	NA

one detector is very low and the other is high. The experts thus disagree and depending on whether the false alarms or the probability of detection is considered the expert with the highest or lowest threshold value should be chosen. In the non-symmetric, concordant case the results could be improved by fusion. A possible fusion strategy could then exist of two successive steps. The first (fastest or cheapest) espert is used for the first classification and depending in the confidence in the results the second expert is used in the second stage.

In the symmetric, concordant case the results can be improved by fusion if the elements $n_{12} > 0$ and $n_{21} > 0$ because this means that the two detector have a similar number of false classifications but the miss-classified elements are not the same: $n_{12} > 0$ and $n_{21} > 0$ are the number of pixels that are classified by one detector as belonging to an edge and by the other detector as belonging to the background. The symmetric concordant case is thus split in two sub-cases, one where the discordant elements are zero and one where they are significantly different from zero. The latter is the ideal case for applying fusion because both experts globally agree but locally can give a different advice. A χ^2 test was used to see whether the discordant elements are larger than the value they would have when the two detectors were independent.

Fig. 5.67 shows the results for the false alarm (left) and detection (right) matrices. The color code is explained below:

McNemar Test	Kappa Test	Consequence for Fusion	Color
Symmetric	Concordant	Fusion useful	Dark Blue if $n_{12} > 0$ and $n_{21} > 0$
		Fusion not useful	Light Blue if $n_{12} \sim 0$ or $n_{21} \sim 0$
	Non-Concordant	Fusion not useful	Red
	Discordant	NA	
Non-Symmetric	Concordant	Fusion may be useful	Green
	Non-Concordant	Choose the best expert	Brown
	Discordant	NA	

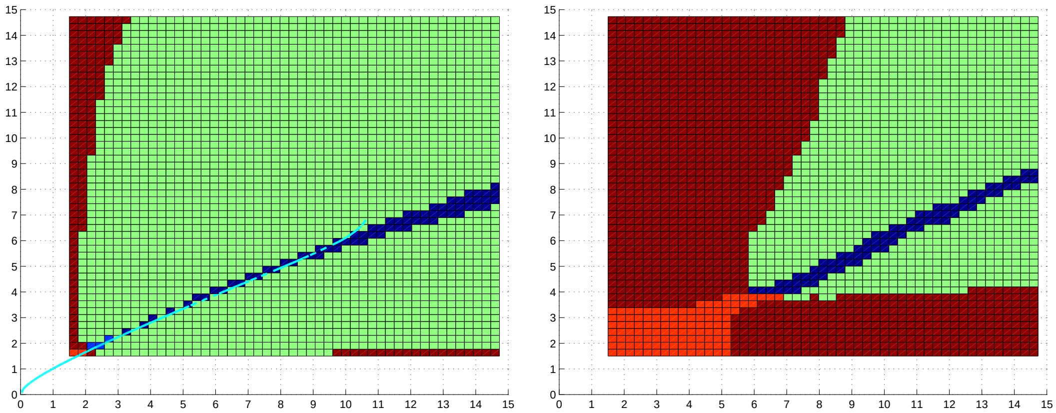


Figure 5.67: Classification of 2D Threshold Space in the 7 Classes

The area corresponding to the symmetric, concordant case with non-zero discordant elements is the most interesting for fusion because the two detectors have globally a similar performance and yet their opinions differ in individual pixels of the test image. The area is linear and the area for probability overlaps with the one for false alarms.

The dashed line in the left figure of fig. 5.67 shows the line of equal probability of false alarms for the two detectors for sampling on a fixed grid and taking the maximum of the result of the 8 orientations of the scanning rectangles. The area that is optimal for fusion coincides partly with the line of equal probability of false alarms in the space of the two thresholds, it deviates at high thresholds.

Our initial idea was to fuse the results of the two detectors by taking into account the relation between the detector's threshold and the probability of false alarms, combining responses that correspond to the same probability of false alarms. The fact that this line does not correspond to the area of optimal fusion, indicates that this first approach was wrong. This was confirmed by the results that were obtained in this manner. We will therefore combine results for sets of thresholds that correspond to the linear area found in fig. 5.67.

The Figure of Merit

Any detection algorithms makes some compromise between the probability of detection and the probability of false alarms. As already mentioned, a ROC curve can be used to see the evolution of the compromise when the threshold of the algorithm is varied. An alternative way is to define a figure of merit for detection. Such a figure of merit was defined in [49, 50] as:

$$FOM' = \frac{N_{dt}}{\alpha N_{dt} + \beta N_{fa} + N_{tt}} \tag{5.42}$$

with N_{dt} the number of detected targets or correctly detected edge points, N_{tt} the number of true edge points, N_{fa} the number of false detections. This is normalised to probabilities by dividing the number of detections by the number of true edge points and dividing the number of false alarms by the total number of pixels within the test image that do not fall within a specified band around the true edges:

$$FOM = \frac{P_{dt}}{\alpha P_{dt} + \beta P_{fa}} \tag{5.43}$$

The parameters α and β are such that $\alpha + \beta = 1$ and they allow to vary the importance that is attached to false alarms or detected points. When $\alpha = high$, $\beta = low$ the most importance is attached on probability of detection; i.e. false alarms are only weakly penalised. The FOM can be calculated for a detector as a function of the threshold. The maximum of the FOM will shift when the parameters α and β are varied. For a given set of these parameters, the maximum of the FOM indicates the optimal threshold for the detector. If two detectors are available, the FOM can be calculated for each of them and the product of the two FOMs gives the combined figure of merit that would be achieved if the results of the two detectors were fused using an "AND" operator (the sum of the FOMs gives the result for an "OR" operator).

The maximum of the 2-D FOM gives the optimal combination of the two thresholds for the corresponding operator.

In fig. 5.68 the results of the statistical analysis for Pd and FA were combined, showing that the linear regions for both the probability of detection and false alarms do coincide. On the figure the contour plot of the figure of merit (for $\alpha = 0.5, \beta = 0.5$) has been superimposed. The central ellipse delimits the maximum of the FOM. In the left figure the results are shown for the first test region, on the right the results are shown for the second test region.

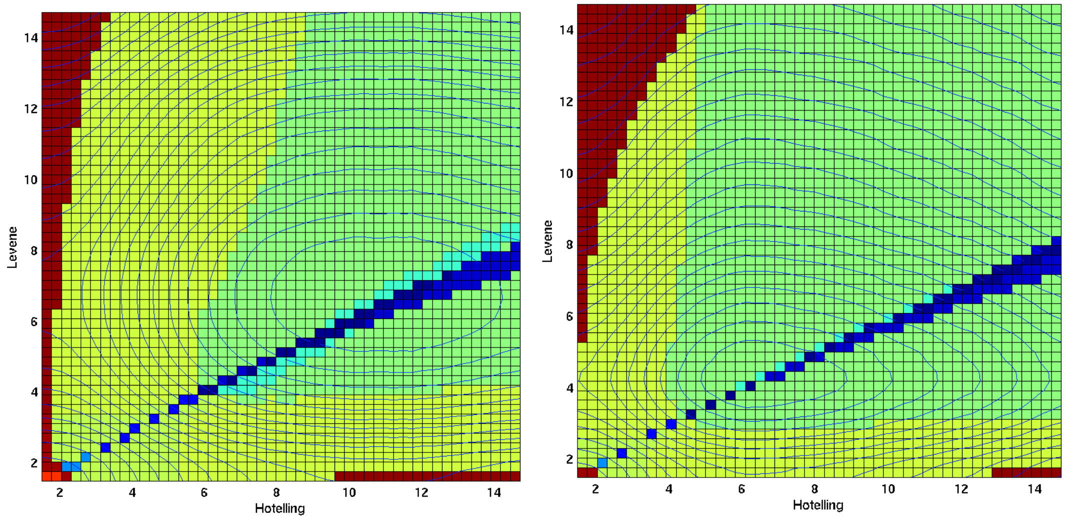


Figure 5.68: Contour plot of FOM ($\alpha = 0.5, \beta = 0.5$) superimposed on the combined results of the classification of the 2D threshold region for probability of detection and false alarms.

By calculating the FOM for different values of α and β and different, ground truthed, test images we have observed that:

- the maximum of the FOMs coincides with a point of the "optimal fusion line" that was found by combining the two statistical tests.

- when changing the values of α and β, the maximum of the 2D FOM moves along the optimal fusion line

- the position of the maximum is the same for the AND and OR operators

- for a given set of values of α and β the maximum of the 2D FOM is different for different test images

The second observation is the most important as it implies that the optimal set of thresholds for a given choice of α and β is the same for AND type fusion operators as for OR type fusion operators. As most fusion operators are a compromise between the two, the same set of thresholds is probably optimal for any fusion operator. The fact that the position of the maximum of the FOM depends on the test image is not surprising because the local content of the image determines the false alarm rate and the probability of detection.

The Fusion Strategy

Several methods to combine the results of two or more detectors are described in literature. In [47] an overview of fusion operators is presented. To apply the fusion methods, the results of the two detectors are first rescaled around a "working point". The working point is a set of thresholds along the linear region of optimal fusion found in the previous paragraphs. On the rescaled image the different information fusion operators can be applied. In fig. 5.69 the result of applying the minimum (in red) and the maximum (yellow) operator on a part of a SAR image is shown (note that the red lines are superimposed over the yellow lines because the pixels selected by the minimum operator are a subset of those selected by the maximum operator). Most fusion operators result in a compromise between the minimum and the maximum operator. From the figure it appears that the result of these two extreme operators is already very similar.

The non-associative sum operator[47] seems to give the subjectively best results (see fig. 5.70). This is consistent with the results that were found for the fusion of uni-variate detectors (over the different polarisations).

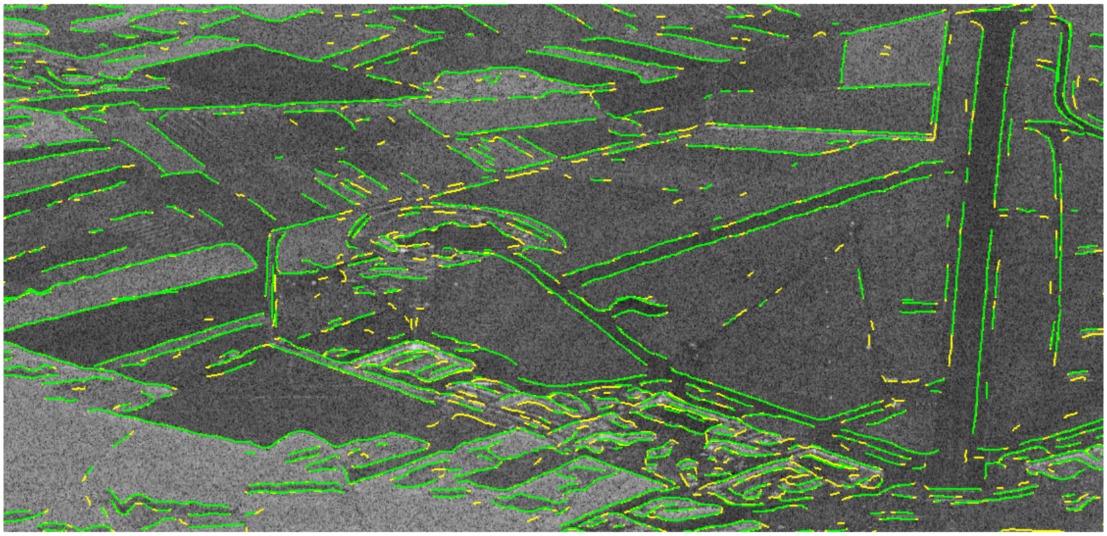


Figure 5.69: Fused Results of Levene and Hotellings Test (Yellow: MAX operator, Red: MIN operator)

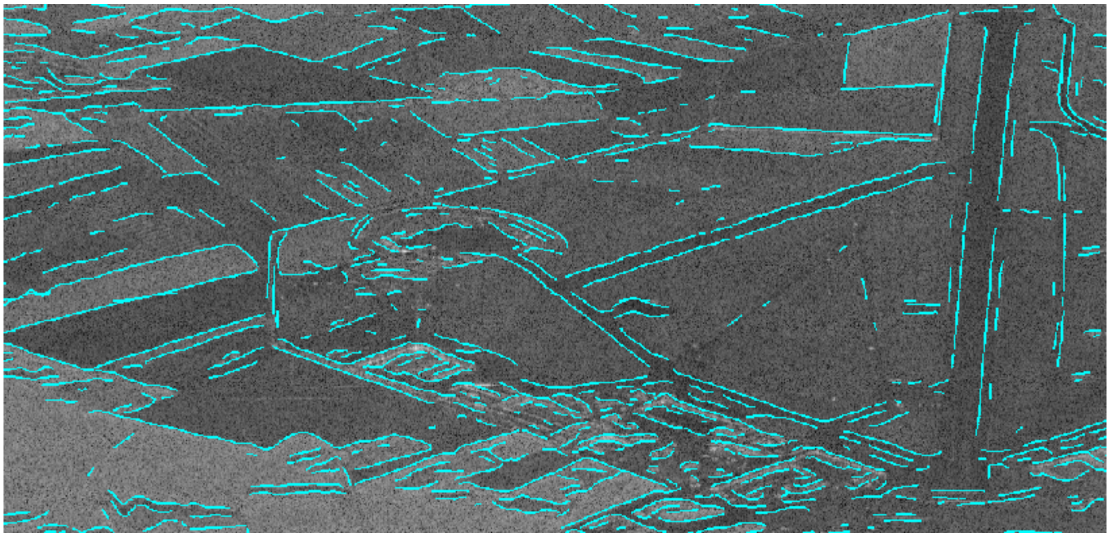


Figure 5.70: Fused Results of Levene and Hotellings Test using a Non-Associative Sum

5.8 Detection of Communication Lines

The aim of the contour detectors that were described in this chapter was to find edges between large regions. This is one of the features that can be used for registration of different images. Edges between large fields are good candidate features for refining a registration between different SAR images because they represent objects with a low 3D structure (see further). In order to find an accurate registration between a topographic map and a SAR image, it is useful to find objects with a low 3D structure in the SAR image that are also visible on maps. Good candidates are communication lines, i.e. roads, railways and waterways. Because their position detected in the SAR image will be used for registration, the localisation should be good and the detection reliable.

In order to detect communication lines a bar detector (narrow lines) needs to be developed. In [32, 30] a method for building a bar detector is presented. The idea is similar to the one that is used for the edge detectors: In order to determine whether a road passes through a given point, in a given direction, a set of rectangles is defined around the point and some statistics are determined in each of the rectangles. The decision is based on a difference between the values of the statistics obtained in the different rectangles. This time three rectangles are used. The middle one is centered on the current point and is narrower than the two others. In fig. 5.71 the principle is represented schematically.

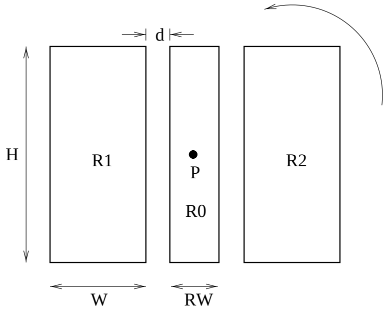


Figure 5.71: General principle for line detection in SAR images

In order to detect whether a line in a given orientation passes through the point P the statistics of the two outer rectangles are compared with those in the middle one. The response of the detector is the response corresponding to the smallest difference in statistics. This allows to distinguish edges from lines and automatically copes with the problem of having a different contrast on both sides of the road. The width of the central rectangle has to be chosen such that it corresponds to the possible widths of the lines to be detected.

When searching for roads on an image of known spatial resolution sensible boundaries for possible road widths can be set. Furthermore, if the image has already been roughly registered with a map, the information from the map can be used to set these boundaries. The gaps between the middle and outer rectangles not only allow to reduce the correlation between the three rectangles, but also builds-in a tolerance for the possible widths of the roads.

The surface of road is very smooth compared to the wavelength of the radar. The

Rayleigh criterion considers a surface to be smooth if its height variations ΔH fulfill:

$$\Delta H << \frac{\lambda}{8cos(\theta)}, \tag{5.44}$$

with λ the wavelength of the radar and θ the local incidence angle. For the E-SAR system the incidence angle varies from 25^o to 57^o over the swath. This yields respectively 3 and 5 cm for the Rayleigh condition in an L-Band image ($\lambda = 23cm$). The surface of the road will thus behave as a mirror. In many cases roads will therefore appear very dark in the SAR image. However, due to the direct environment of the road and in particular the presence of shrubs, trees, ditches boarding the road, double bounce reflections can occur and the road will appear very bright [30]. This effect only occurs for roads that are almost perpendicular to the viewing direction of the radar.

Rivers have similar characteristics as roads. They also behave mainly as smooth surfaces. Railways appear as very bright lines in a SAR images. This is due to double bounce reflections and the presence of metallic structures.

In order to detect the communication lines we will apply the same methods as for edge detection to verify the presence of an edge between both of the outer rectangles and the central one. Because this time we would like to detect narrow structures as opposed to edges between large uniform regions, it is necessary to use as much pixels within each rectangle as possible.

For the uni-variate edge detector based on the Student test, the percentual variance of the detector's response over different regions does not vary with sampling ratio. Furthermore the average threshold at full resolution is the same for vertical and horizontal windows and it corresponds to the theoretical threshold multiplied by the correction factor. The Student test based detector is therefore a good candidate for detecting line structures.

Unfortunately neither of the multi-variate detectors has the same threshold at full resolution for horizontal and vertical scanning windows. Using them would thus require a different threshold for each orientation of the scanning windows.

In fig. 5.72 the results of the Student-test based bar detector are shown (parameters used: $RW = 3, d = 2, W = 20, L = 30$). The original image is shown in green and the bright and dark bars are overlaid in respectively red and blue. Dark and bright bars are detected by imposing the sign of the difference in means between the central and outer scanning rectangles. Note that for the large buildings of the airfield one does find a bright bar corresponding to the wall facing the radar and a dark bar at the opposite side of the building.

For detecting roads, a threshold is applied to the outcome of the detector of dark bars and the results are vectorised. The results of this are shown in fig. 5.73. Only some segments of the roads are detected and shadows (e.g. from forests) produce false alarms.

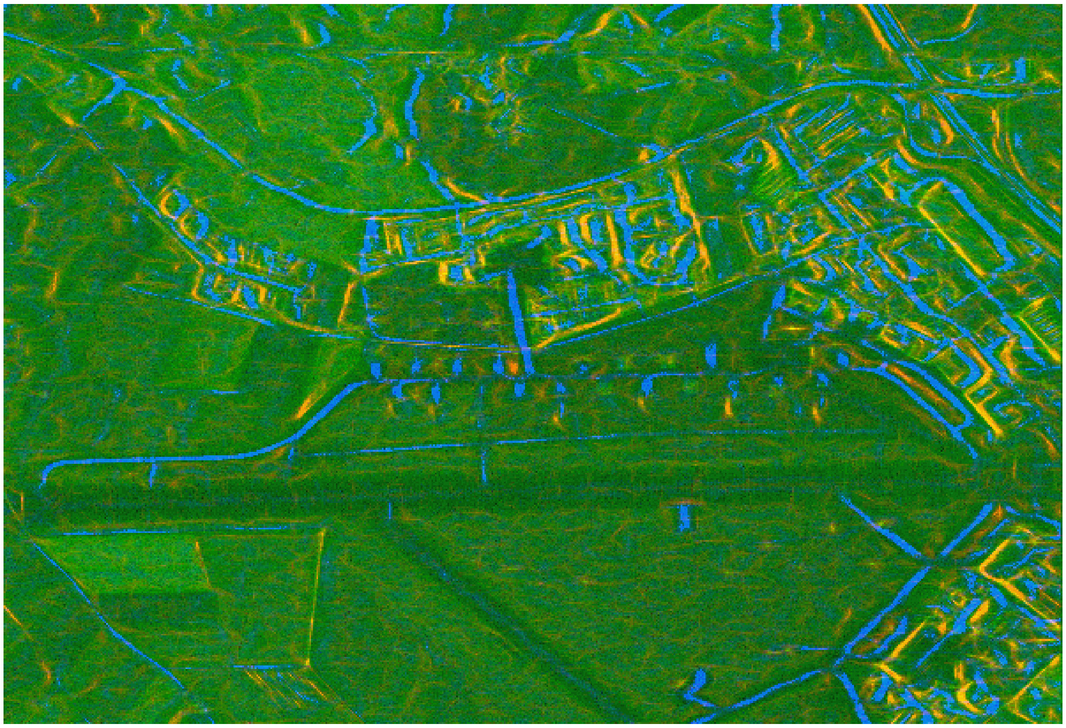


Figure 5.72: Results of the Student-test based bar detector (blue=dark bars, red=bright bars)

Figure 5.73: Vectorised results of the Student-test based detector of communication lines