

A Ship Detection Algorithm Based on Truncated Statistics

Ding Tao, Stian Normann Anfinssen, and Camilla Brekke

Department of Physics and Technology, University of Tromsø – The Arctic University of Norway

Email: stian.normann.anfinssen(at)uit.no, Tel: +47 776 45173

Abstract

A new constant false alarm rate detector is proposed for ship detection in single-look and multilook intensity synthetic aperture radar images. The method is aimed at multiple target situations where the sea clutter statistics are estimated from a sample which is potentially contaminated by targets. It uses truncation to exclude outliers from the sample and truncated statistics to analyse the truncated sample in a statistically rigorous manner. Experiments show that the detector performs on par with state-of-the-art methods at lower computational cost, has excellent false alarm regulation properties, and can estimate sea clutter statistics from a window centered at the cell under test.

1 Introduction

Ship detection is an important application of spaceborne synthetic aperture radar (SAR) images, and has been turned into operational services around the world. The most common method in use is the constant false alarm rate (CFAR) detector, which has been implemented in many versions. The principle behind CFAR detectors is that a statistical model for the background sea clutter is hypothesised and its parameters estimated locally in the images. Ships are detected as outliers that stand out from the statistical background model by their high intensity, without any further modelling of the target pixels. A detection threshold is selected such that the decision rule is associated with a specified false alarm rate (FAR), i.e., probability of producing false positive decisions. As long as the model assumption is satisfied and the parameter estimation is accurate enough, the observed FAR will in practice be approximately constant, and should also match the specified FAR.

The simplest CFAR algorithm is the cell averaging CFAR (CA-CFAR) detector [1]. It assumes that the cell under test (CUT) is surrounded by spatially homogenous sea clutter, apart from the potential sidelobes from a target within the CUT. The background is characterised by the mean intensity computed in an estimation sample, from which model parameters can be estimated. To avoid that energy from sidelobes interferes with the estimation of sea clutter statistics, a guard area is introduced, such that the estimation sample is confined to the corners shown as shaded fields in **Figure 1(a)**. An alternative method is shown in **Figure 1(b)**, where an estimation window centered at the CUT is chosen. The second approach collects samples closer to the CUT and is therefore less affected by nonstationary clutter. It will be explained how the method proposed in this paper may use a centered estimation window without being disturbed by sidelobes.

Alternative CFAR detectors have been designed that handle deviations from the homogeneous clutter assumption. The proposed algorithm is intended for situations with high target density, e.g., ship detection around busy

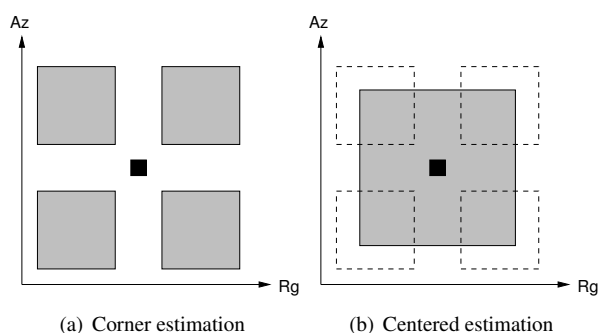


Figure 1: Alternative ways of selecting the estimation window for sea clutter statistics. The CUT is marked in black and the estimation sample is shaded.

shipping lines and crowded harbours. In such cases, there may be one or more ships inside the estimation sample. This will lead to overestimation of the clutter intensity, a raised detection threshold, and thus masking of targets and lower detection rates. Previous attempts to handle this situation include the smallest-of CFAR (SO-CFAR) detector [2] and the order statistics CFAR (OS-CFAR) detector [3]. The SO-CFAR algorithm subsets the estimation sample spatially, computes the mean intensity in all subsets, and uses the lowest value as the estimate. The OS-CFAR algorithm rank orders the intensities in the estimation window, uses the value with a certain rank to represent the sample, and relates this so-called order statistic to model parameters and the detection threshold. Another tool for multiple target situations is iterative censoring (IC), which was proposed by Barbooy et al. [0] and revisited in [0]. The detection result defines an outlier map. All outliers are excluded from the estimation sample, before parameter estimation, threshold computation and hypothesis testing is reiterated. The algorithm converges when the outlier map stabilises. IC is a generic scheme which can be applied to any CFAR algorithm, such as the CA-CFAR and OS-CFAR detectors, that are respectively turned into the ICCA-CFAR detector and the ICOS-CFAR detector. The iterations have a relatively high computational cost, but the good performance po-

sition these methods as state-of-the-art.

The new CFAR algorithm we propose is motivated by some critical remarks about how the IC scheme is implemented: The censoring removes outliers from the estimation sample, including both targets and high intensity sea clutter spikes (false alarms), that cannot be distinguished. Hence, the estimation sample may be deprived of observations in the upper tail of the sea clutter distribution, which biases the estimates of mean intensity (or other statistics) towards values that are too low. The latter is a consequence of using standard estimators designed for uncensored samples, which is how the IC scheme has been realised [0, 0]. This is also our entry point for making improvements.

A reduced estimation sample can be treated in a statistically rigorous manner, if we devise an estimator which accounts for the data reduction. When outliers are removed, it is unknown how many of the outliers that are true targets and how many that are sea clutter spikes. Now the distinct definitions of censoring and truncation becomes important: the number of censored data points is known, whereas the number of truncated data points is unknown. Thus, the reduced sample must be regarded as truncated and handled according to the theory of truncated statistics.

Section 2 introduces truncated statistics and presents the truncated statistics CFAR (TS-CFAR) algorithm, including parameter estimators for exponentially distributed single-look intensity (SLI) data and gamma distributed multilook intensity (MLI) data. Section 3 provides simulation results of the detector operating characteristics and false alarm regulation results, where the TS-CFAR detector is compared with alternative detectors. Performance results on modified real data from Radarsat-2 are also given. Section 4 gives our conclusions.

2 TS-CFAR Algorithm

2.1 Truncated Statistics

Suppose we have a random variable, X , which follows statistical distribution with probability density function (pdf) $p_X(x)$ and cumulative distribution function (cdf) $P_X(x)$. Let \tilde{X} be the truncated version of X after applying a threshold t to the upper tail of $p_X(x)$, where t is called the truncation depth. The pdf of the right truncated distribution can then be defined as

$$p_{\tilde{X}}(x; t) = p_X(x|X < t) = \frac{p_X(x)}{P_X(t)}. \quad (1)$$

Thus, $p_{\tilde{X}}(x; t)$ maintains the shape of $p_X(x)$, but over a limited domain, and the normalization by $P_X(t)$ makes $p_{\tilde{X}}(x; t)$ integrate to one. The truncation depth is a user specified value, which should be set to ensure that all possible targets are excluded. A practical solution is to remove a fraction R_t of the estimation sample, where R_t is called the truncation ratio, and define the truncation depth as the smallest of the truncated intensities.

2.2 Parameter Estimation

We will now derive parameter estimators for the mean value in a truncated exponential distribution and a truncated gamma distribution. These are the assumed models for truncated single-polarisation image data on SLI and MLI format, respectively.

2.2.1 Exponential Distribution

The SLI measurements are represented by the random variable $X \geq 0$, that are assumed to follow the exponential distribution with pdf

$$p_X(x) = \frac{1}{\mu} e^{-x/\mu}; \quad x \geq 0 \quad (2)$$

and cdf

$$P_X(x) = 1 - e^{-x/\mu} \quad (3)$$

where $\mu > 0$ is the mean value. Let $0 \leq \tilde{X} < t$ be a SLI measurement which is right truncated with truncation depth t . The pdf of \tilde{X} becomes

$$p_{\tilde{X}}(x; t) = \frac{1}{\mu} \frac{e^{-1/\mu}}{1 - e^{-t/\mu}}; \quad 0 \leq x < t. \quad (4)$$

The only parameter which needs to be estimated is the mean value. Assuming that we have a sample $\{x_i\}_{i=1}^n$ of n independent SLI measurements distributed according to Equation (4), the maximum likelihood estimator (MLE) for μ can be obtained as [0]

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i + \frac{t}{e^{t/\hat{\mu}} - 1} \quad (5)$$

which must be solved numerically. We note that the first term equals the MLE for the standard exponential distribution, while the second term corrects for the effect of the truncation operation.

2.2.2 Gamma Distribution

An L -look MLI measurement $X_L \geq 0$ can be computed as the sample mean of a set $\{x_i\}_{i=1}^L$ of independent and exponentially distributed SLI measurements with mean value μ . It follows that X_L is gamma distributed with pdf

$$p_{X_L}(x) = \frac{1}{\Gamma(L)} \frac{L}{\mu} \left(\frac{Lx}{\mu} \right)^{L-1} e^{-\frac{Lx}{\mu}} \quad (6)$$

where $\Gamma(\cdot)$ is the gamma function, and cdf

$$P_{X_L}(x) = \frac{\gamma(L, Lx/\mu)}{\Gamma(L)} \quad (7)$$

where $\gamma(\cdot, \cdot)$ is the lower incomplete gamma function. The pdf of the truncated MLI, with domain $0 \leq \tilde{X}_L < t$, is then derived as

$$p_{\tilde{X}_L}(x) = \frac{1}{\gamma(L, Lt/\mu)} \frac{L}{\mu} \left(\frac{Lx}{\mu} \right)^{L-1} e^{-\frac{Lx}{\mu}}. \quad (8)$$

Due to correlation between the SLI samples used to compute X_L , the value of L is conventionally reduced from the nominal number of averaged samples to a noninteger

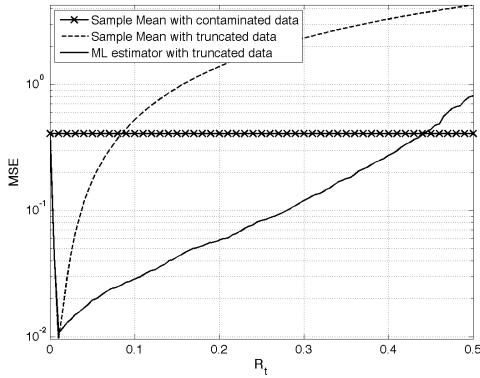


Figure 2: MSE versus R_t for three estimators of μ applied to simulated data with $R_c = 1\%$.

value known in SAR terminology as the equivalent number of looks. This value can be estimated from data and is assumed to be a known image constant. Thus, only the mean value μ remains to be estimated. A MLE for the truncated gamma distribution exists, but the probability that it provides a solution is less than one [0]. We therefore resort to the method-of-moments (MoM) estimator, which solves the moment equation

$$E\{(\tilde{X}_L)^m\} = \left(\frac{\mu}{L}\right)^m \frac{\gamma(L+m, Lt/\mu)}{\gamma(L, Lt/\mu)} \quad (9)$$

at moment-order $m = 1$ after the expectation $E\{\tilde{X}_L\}$ has been replaced by a sample mean of \tilde{X}_L . This solution $\hat{\mu}$ must be sought numerically.

2.2.3 Estimator Performance

To assess estimator performance, we have simulated exponentially distributed sea clutter, where a fraction R_c of the pixels have been replaced by high intensity samples that resemble targets. The target samples are drawn from a uniform distribution whose support is 0.8 to 5 times the maximum intensity in the simulated sea clutter sample. We refer to R_c as the contamination ratio, and set it to $R_c = 1\%$ in the experiment.

Figure 2 shows the performance of the proposed MLE for the truncated exponential distribution (solid line), as measured by the mean squared error: $MSE = E\{(\hat{\mu} - \mu)^2\}$ for a range of different R_t values. It is compared to the MLE for the standard exponential distribution, which is the simple sample mean estimator (SME). The SME is run both on the truncated sample (dashed line) and on the complete sample (line with crosses). The truncation ratio obviously has no effect in the latter case and the MSE is constant across R_t .

We observe that the MLE for the truncated model is consistently better than the SME applied to truncated data, and is only surpassed by the SME applied to the complete data for values of R_t that are irrelevant in practice. The estimators that apply truncation show an increase of MSE with R_t , as would be expected when data is progressively discarded. However, the gain in efficiently removing outliers is evident, at least for the proposed MLE

estimator, which is derived from the correct model for truncated data. Results for the MLI case with the MoM estimator based on the truncated gamma distribution are similar, and will be presented in a journal paper.

3 Detector Performance

This section presents a quantitative comparison of false alarm regulation properties and detector operating characteristics (DOC). We here compare the classical CA-CFAR and OS-CFAR detectors, their IC implementations: ICCA-CFAR and ICOS-CFAR, and finally the proposed TS-CFAR detector.

3.1 False Alarm Regulation

The false alarm regulation property is the ability of a detector to produce an observed (or actual) false alarm rate (FAR) which matches the specified FAR. The property is assessed by measuring the ratio: P_{fa}/P_{FA} , where P_{fa} is the observed FAR and P_{FA} the specified one. The ratio is given in dB, such that values around zero are desired.

Figure 3 displays plots of P_{fa}/P_{FA} versus P_{FA} , as measured under different degrees of contamination, ranging from $R_c = 1\%$ to 20%. Since we compare many P_{fa}/P_{FA} values across different scales, the results are presented in three plots: The upper panel compares CA-CFAR and ICCA-CFAR; The middle panel compares OS-CFAR and ICOS-CFAR; and the lower panel assesses only TS-CFAR. The results show that CA-CFAR and OS-CFAR exert bad false alarm regulation. Their implementation with the IC scheme greatly improves the regulation. However, the plots show that TS-CFAR is superior by far, and particularly around the low P_{FA} values commonly used in operational systems ($P_{FA} \sim 10^{-5}$).

3.2 Detector Operating Characteristic

Figure 4 shows traditional DOC curves where the detection rate, P_D , is plotted against P_{FA} . Various R_c levels are considered and the TS-CFAR detector is implemented with different values of R_t .

The results show that the CA-CFAR detector does not handle the multiple target case, the OS-CFAR detector performs much better, but not nearly as well as the ICCA-CFAR, ICOS-CFAR, and TS-CFAR detectors, that we claim represent state-of-the-art. Remark that it would favour the TS-CFAR detector if P_D was assessed against P_{fa} instead of P_{FA} , since the competitors produce much more false alarms than specified at operational FAR levels. Further note that the TS-CFAR avoids iterations and therefore saves computation time.

4 Conclusions

We have presented a new CFAR algorithm based on truncated statistics, which has a number of advantages. It performs on par with iterative censoring algorithms, while

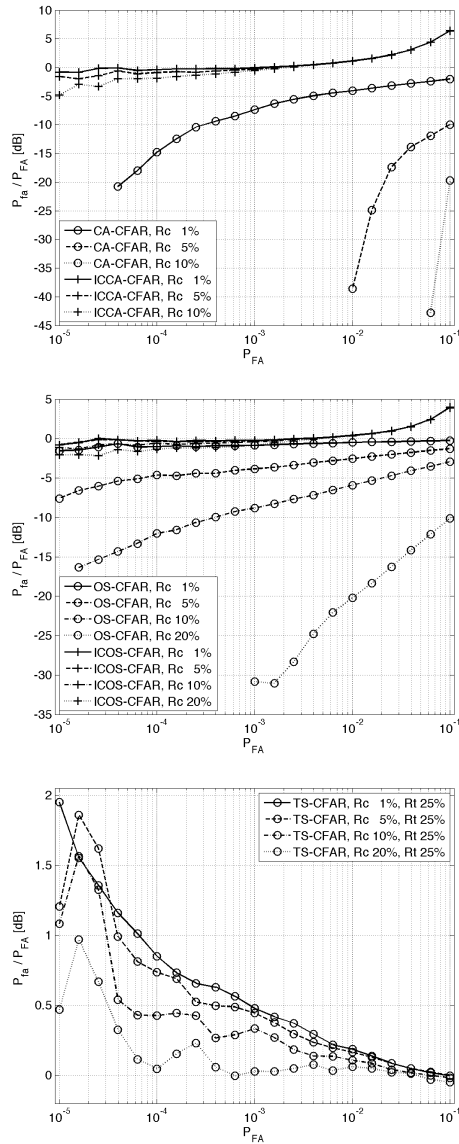


Figure 3: Comparison of false alarm regulation properties where the ratio between observed and specified FAR (in dB), P_{fa}/P_{FA} , is plotted against P_{FA} .

avoiding the iterations. It regulates the FAR exceptionally well. Due to the truncation, it can be implemented with a sliding estimation window centered at the CUT, thus collecting the sample in a more confined area.

References

- [1] H.M. Finn and R.S. Johnson: *Adaptive detection mode with threshold control as a function of spatially sampled clutter-level estimates*, RCA Review, vol. 29, pp. 414–464, 1968.
- [2] G. Trunk: *Range resolution of targets using automatic detectors*, IEEE Trans. Aerosp. Electron. Syst., vol. 16, no. 1, pp. 750–755, 1978.
- [3] H. Rohling: *Radar CFAR thresholding in clutter and multiple-target situations*, IEEE Trans. Aerosp. Electron. Syst., vol. 19, no. 4, pp. 608–621, 1983.

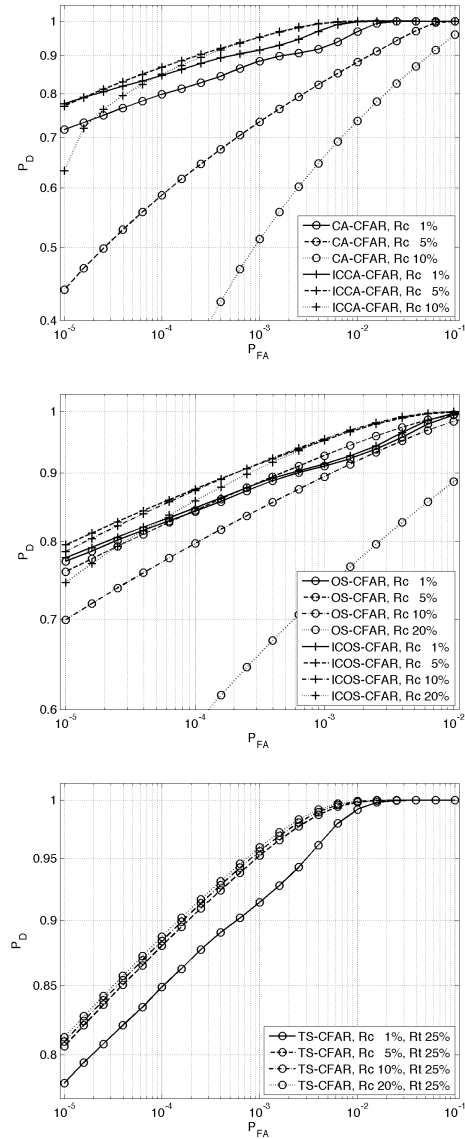


Figure 4: Comparison of detector operating characteristic curves that plot the detection rate, P_D , against the specified FAR, P_{FA} .

B. Barbo, A. Lomes, and E. Perkalski: *Cell-averaging CFAR for multiple-target situations*, IEE Proc. F Commun. Radar Signal Process., vol. 133, no. 2, pp. 176–186, 1986.

Y. Cui, G. Zhou, J. Yang, and Y. Yamaguchi: *On the iterative censoring for target detection in SAR images*, IEEE Geosci. Remote Sens. Lett., vol. 8, no. 4, pp. 641–645, 2011.

N. Balakrishnan and B.K. Sinha: *Order Statistics and their Properties*. Ch. 3 in *The Exponential Distribution: Theory, Methods and Applications*, N. Balakrishnan and A.P. Basu (eds.), Gordon and Breach Science Publishers, 1995

L.M. Hegde and R.C. Dahiya: *Estimation of the parameters of a truncated gamma distribution*, Commun. Statist. – Theory Meth., vol. 18, no. 2, pp. 561–577, 1989.