UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

IMPROVED STOCHASTIC DETECTION ALGORITHMS WITH APPLICATIONS IN RADAR AND LADAR BASED ROBOTIC MAPPING

TESIS PARA OPTAR AL GRADO DE DOCTOR EN INGENIERÍA ELÉCTRICA

DANIEL VICENTE LÜHR SIERRA

PROFESOR GUÍA:

Ph.D. MARTIN ADAMS

MIEMBROS DE LA COMISIÓN:
DR. ING. JAVIER RUIZ DEL SOLAR SAN MARTÍN
PH.D. MIGUEL TORRES TORRITI
DR. DANIEL CLARK

Este trabajo ha sido parcialmente financiado por el Programa de Becas para estudios de Doctorado año 2010 de CONICYT

SANTIAGO, CHILE
2018

RESUMEN DE LA TESIS PARA OPTAR AL GRADO
DE DOCTOR EN: Ingeniería Eléctrica
POR: Daniel Vicente Lühr Sierra
FECHA: 2018-01-12
PROFESOR GUÍA: Martin Adams

Improved stochastic detection algorithms with applications in radar and ladar based robotic mapping

In this work a set of robust tools for noise reduction and feature detection in radar and ladar environment data for outdoor robotic applications have been developed.

One of these tools is a new noise reduction technique for radar data which combines the well known spectral noise subtraction with the standard constant false alarm rate (CFAR) detectors and binary integration. The resulting method exhibited a relatively low time complexity compared to other state-of-the art noise reduction techniques, the Wiener filter and Wavelet denoising, while retaining a higher Signal-to-Noise ratio. The method was tested with real data from a local park captured with a scanning radar mounted on a robot platform and with SAR data available from NASA/JPL UAVSAR missions.

The second component of this framework is the extension of the standard CFAR detectors used in radar data to be used with 3D ladar point cloud data. More generally, these extensions can be used with any kind of 3D point cloud data which comply with the stochastic CFAR assumptions. The extended CFAR detectors are capable of processing a 2D projected version of the 3D data or they can work directly on the 3D point cloud. The main modifications to the original methods include making the CFAR window size parameter an adaptive one and adding the capability to work with sparse data, in contrast to dense data which is what the original methods expect. The extended CFAR detectors show a more robust performance than other methods when the point cloud data contains high noise and clutter rates. The output of these detector applied to ladar 3D data could then be used for algorithms requiring high accuracy in the detection, for instance, the Iterative Closest Point (ICP) registration method.

To God for the talents and opportunities with which I have been gifted...
To Grace, my wife, for her long lasting patience...
To my parents for their support and guidance...
To my sister, my brother-in-lay and their daughter, my godchild, for being a source of inspiration and joy...
To my friends for asking me every other day when I am going to finish (and for motivating me to keep going).

Acknowledgements

I would like to thank my advisor Dr. Martin Adams, for his patience. I would also like to thank Dr. Javier Ruiz del Solar, for believing in me, for motivating (and convincing) me to start this adventure and for his support all along. I would like to thank Señora Eliana and now Milena for all the paperwork "miracles", for making the impossible to solve the unsolvable and succeeding for the good of all of us (the postgraduate students). I would like to thank my fellow lab mates from both the DIE's 3rd floor and the AMTC's 1st floor for their time to discuss, to drink a cup a coffee and to jointly complain about our sufferings in the program. I would like to thank Dr. Andreas Nüchter and his team for giving me and our team the opportunity to carry out interesting projects together. I would like thank Daniel and Jeremy for organising the great FiSSt summer school at Herriot-Watt University, and for supporting our participation.

Also, I would like to thank CONICYT for their financial support in the form of the Scholarship "Beca Doctorado Nacional 2010" and the "Programa de Cooperación Científica Internacional" which were fundamental for the partial funding of my work. I would like to thank the UChile's AMTC, which has brought great, interesting challenges from the real world to the Academia, as well as the resources needed to face them. Finally, I would like to thank UACh for partially financing my work.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Mobile robots and autonomous vehicles navigating in outdoor, usually unstructured, environments often need to localize themselves and map uncharted terrain in order to perform their assigned tasks. In order to solve the localization and mapping problem, it is necessary to perceive the environment, and store a representation of that environment. To accomplish this, a robot relies on sensing technologies which retrieve data from the environment. A brief list of sensor technologies used in robotics includes:

- Infrared range sensors
- Sonar range sensors
- Laser range sensors
- Radar sensors
- Video Cameras

Cameras are very rich in information, while laser range sensors provide robust and accurate range measurements. Laser range sensors are used to build 3D laser scanners or «ladars», which are capable of making 3D-representations of targets. In general, electromagnetic waves, such as those used by radar sensors, can be reflected, absorbed or transmitted through materials receiving this radiation. How much is reflected, absorbed or transmitted depends on the nature of the material and the wavelength of the transmitted signal. Because of this, certain radar systems can be used to detect multiple objects within a single line-of-sight. They are capable of penetrating the dust, fog, rain and snow conditions usually found in mining,

construction, agricultural and other environments, which often interfere with measurements from laser range sensors and cameras. These are the main advantages of radar over other sensors in field-robotics applications. However, radar data is corrupted by noise from spurious measurements as well as clutter[1]. In some situations, before a robot can use radar data to identify the environment, it has to be pre-processed.

In the literature, the environment has been represented in fundamentally different ways. In particular, occupancy grid, Feature Based (FB) and topological maps, have been applied in autonomous mapping research. Of these, occupancy grids and FB maps have emerged as the most popular means of environmental representation [MVAV11], and are typically modelled in a probabilistic manner to account for sensor and vehicle motion uncertainty.

FB mapping approaches offer the advantage of compressing sensor data into features (such as point clusters, circles, lines, corners etc.). The feature map representation has gained widespread popularity, particularly in the outdoor robotics domain due to its ability to estimate large scale maps and provide correctional information for simultaneous vehicle pose estimation.

FB approaches can be computationally appealing since few features need to be detected and updated per sensor scan and, feature-to-feature and feature-to-vehicle correlations can be maintained. They fundamentally differ from their grid based counterparts, in that they are required to estimate the location of an initially unknown number of objects, represented as features. Hence, the number of features and their locations, which fully describe the environment, are typically represented by *varying* the size of a vector.

Once a robust method of landmark detection has been established, its integration into a robotic navigation and mapping framework is necessary. The Simultaneous Localization and Mapping (SLAM) in robotics copes with the problem of building a map of an unknown environment by a mobile robot while, at the same time, navigating the environment using the map [RB03].

Realistic feature detection algorithms produce false alarms (as well as missed detections), and estimating the true number of features is therefore central to the SLAM problem. The stochastic information available from adaptive detection algorithms (such as *Constant False Alarm Rate* processors), namely the probability of detection and probability of false alarm, is not always considered in current detection techniques used to detect features of the environment.

Another critical aspect arises when measuring the accuracy of the position and number of the detected features. Most metrics cannot encapsulate an error in estimating the number of features and therefore are not suitable for the evaluation of FB estimates [MVAV11].

[1] *clutter*: unwanted echoes typically returned from ground, sea, rain, atmospheric turbulence, etc.

1.2 Research proposal

1.2.1 Proposal

In this thesis work, a set of tools for processing environment data captured with radar or ladar sensors will be proposed. The final aim of all these tools is to improve target detection in such data. Also, they are all based on the well established family of stochastic detection algorithms, collectively known as *Constant False Alarm Rate (CFAR) processors*, which are mainly found in radar applications. First, a noise reduction method for radar data is developed. It is based on stochastic data integration of multiple scans combined with the spectral noise subtraction. The new method exhibits a low computational complexity when compared to standard noise reduction techniques and presents a higher performance in terms of reduced noise. The output of this algorithm can then be used as the input for a detection stage in order to find radar targets in the data. Then, classical radar-based stochastic CFAR detection algorithms are extended to new versions which can be used with 3D ladar data, or any other point cloud data, as feature extraction algorithms, previous to a registration process. These extensions where designed to work with 3D data projections in 2D or directly with the 3D point cloud representation.

1.2.2 Hypothesis

It is possible to implement more robust methods for detecting features in complex environments using radar and ladar sensors:

- The detection algorithm will be improved by using a noise reduction method which yields a probability of detection higher than a *Constant False Alarm Rate* processor for a given probability of false alarm.
- The *Constant False Alarm Rate* (CFAR) detection algorithms traditionally used with dense radar data can be modified to work directly with sparse 3D point cloud data and indirectly with 2D projections of the point cloud data by making some of their parameters adaptive, namely the reference and guard cells window parameters.

By using all these individual improved components, the resulting set of tools should present an overall better performance when compared to a standard implementation by means of a second order Wasserstein metric to jointly assess errors in target locations and number of estimates.

1.2.3 General and specific goals

The main goal of this doctoral work is to develop a set of tools, which can be used on radar or ladar data, based on robust noise reduction methods and detection algorithms. Noise reduction methods increase detection probability at a desired false alarm rate.

Specific objectives of this work are:

- Develop, implement and analyze a radar data noise reduction method based on stochastic Binary Integration and Spectral Noise Subtraction.
- Extend classical radar CFAR detection methods for use with 3D data projected onto 2D images.
- Extend classical radar CFAR detection methods for direct application on 3D point cloud data.
- Test the performance of the CFAR extensions in both simulated and real data and compare it to the performance of other state-of-the art detection methods.
- Create experimental radar data sets.

1.3 Thesis structure

Chapter 2 describes the necessary background concepts: the use of radar in robotics (Section 2.1), radar detection and noise reduction techniques (Section 2.2), Feature detection in 3D environments (Section 2.3), and the Optimal Sub-pattern Assignment map metric (Section 2.4).

The following three chapters correspond to the main parts of this dissertation. Chapter 3 presents the Binary Integration Noise Reduction method. Its development and implementation details are described in Sections 3.1 and 3.2, respectively. The performance of the method is compared with two state-of-the art noise reduction techniques: Wiener filtering and wavelet denoising. The analysis is carried out on experimental radar data obtained with a scanning radar in a park in Santiago, Chile (Section 3.3) and with SAR data courtesy of NASA/JPL-Caltech's UAVSAR mission (Section 3.4). Extended version adapted from standard radar-based CFAR methods for use with 3D data are presented in Chapter 4.

Final conclusions and remarks as well as future work is presented in Chapter 5.

Appendix A and B correspond to the published papers related to this thesis work, and technical specifications of the sensors used in this work are listed in appendix C.

1.4 Contribution

The contributions of this thesis work are:

- Develop and implement a noise reduction scheme for radar data, based on Binary Integration and Spectral Noise Subtraction.
- Compare the performance of the Binary Integration Noise Reduction method with other state-of-the-art denoising techniques.
- Extend and implement CFAR detection algorithms for 3D point cloud data (particularly, ladar data).
- Analyze the performance of the adapted CFAR methods.
- Radar data sets of a park environment and an underground mine tunnel.
- Implement code libraries which can be incorporated in standard data processing tools.

1.4.1 Publications

The following article was published as a result of this research[2]:

- Lühr, D.; Adams, M., "Radar Noise Reduction Based on Binary Integration", *Sensors Journal*, *IEEE* , vol.15, no.2, pp.766,777, Feb. 2015, doi: 10.1109/JSEN.2014.2352295

During this work, a multi-sensor data set was produced in a field trip to an underground mine. The radar data was captured and pre-procesed with some of the tools presented in this thesis. The full data set has been published in the International Journal of Robotics Research:

- Keith Y. K. Leung, Daniel Lühr, Hamidreza Houshiar, Felipe Inostroza, Dorit Borrmann, Martin Adams, Andreas Nüchter, and Javier Ruiz del Solar. "Chilean Underground Mine Dataset". *International Journal of Robotics Research (IJRR).*

There is also a work in progress to produce a paper about the extensions of radar detection algorithms to ladar data. Preliminary information about title and authors:

- Lühr, D.; Adams, M.; Houshiar, H.; Borrmann, D.; Nüchter, A., "CFAR-based Automatic Detection of Ladar Features for 3D Laser Surveying".

[2]The article was one of the 50 most downloaded Sensors Journal papers in January and February 2015.

Chapter 2

State of the Art

2.1 Radar in Robotics

Visibility conditions are often poor in field-robotic scenarios. Day/night cycles change illumination conditions. Weather phenomena such as fog, rain, snow, and hail, as well as dust clouds are usually present in excavation and agricultural sites, and can impede visual perception. Also, smoke compromises visibility in fire emergencies and battlefield operations [Foe02a]. Ladar sensors and video cameras are common visually based sensors affected by these conditions. Millimetre-wave (MMW) radar, on the other hand, provides consistent range measurements in such environments [FCA99]. Radar can, in addition, provide information of distributed and multiple objects in a single reading [Foe02b]. The reliability of MMW radar sensors under snow conditions was demonstrated in [FCA99] by testing a short range MMW sensor in a Polar environment. Snow conditions slightly increased noise, but it was not enough to prevent radar sensing of obstacles.

Radar technology specially developed for robotic applications, like autonomous navigation and mining operation, has produced new sensors and processing software which match the specific requirements of the field. Work was conducted to build a radar system comprising a custom made Frequency Modulated Continuous Wave[1] (FMCW) MMW radar front end with a rotating mirror scanner and client-server visualization software [SBH+02] serving point cloud data and Delaunay triangulation for 2D and 3D visualization. This system proved to facilitate autonomous operation of vehicles or machinery during periods of poor visibility. [WCBS+06b] presented an application of MMW radar to environmental mapping in surface mining with shovel and dragline excavators. The radar system showed a successful performance in terms of accurate measurement, real-time imaging and penetration capabili-

[1]For a description of how FMCW radar works see [Wol11]

ties through dust and water. A low cost MMW radar designed for terrain imaging on board an Unmanned Aerial Vehicle (UAV) was described in [JB08]. This system also exhibited low power consumption and weight. These advances in radar technology have made its application in robotics feasible, but have also shown that progress is still needed in processing radar data to achieve robust and reliable solutions.

An early application of radar to the mapping problem was introduced in [FBW01a]. A three dimensional radar-sensor model for occupancy grids was presented. The model used qualitative rules and a CFAR processor, but it lacked a more rigorous probabilistic formulation in order to improve the results. A first attempt to incorporate particular features observed in radar data is described in [JA05]. This work presented an initial formulation for implementing an augmented state, feature rich FB SLAM formulation with MMW radar, introducing absorption cross section features. A method for predicting scanned MMW radar data, based on the FMCW range measurement technique is explained in [JAMP10]. In this work a power versus range prediction is made taking in consideration the presence of noise and its statistical characterization. This method can be used for the data association problem or to simulate radar scans in the presence of noise.

One of the most recent applications of radar in robotics presented an autonomous surface marine craft incorporating a commercial off-the-shelf X-band marine radar in [MKR$^+$10]. This project was intended as a step towards realizing an autonomous craft capable of performing environmental or security surveillance and reporting a real-time active awareness of the above-water scene. Landmark detection was carried out using image processing algorithms and an RFS SLAM framework based on a PHD filter was implemented.

2.2 Radar Detection

Radio Detection And Ranging (radar) is a system capable of determining range, altitude, direction or speed of objects using electromagnetic waves with frequencies ranging from 3 MHz to 300 GHz. The millimetre wave radar band spans from 40 GHz to 300 GHz offering high reliability in bad weather conditions. MMW radar has a comparatively long range which can enable a vehicle to localize even with sparse features in its proximity [Jos09].

The radar system emits radio waves which could be reflected back by objects in their path. The radar receives these returning echos and process their signal to obtain the desired information (range, speed, etc.). Relevant information is the amount of power reflected by objects as it can be used to distinguish relevant objects from noise. The power P returning

to the receiving antenna is given by the equation:

$$P = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4 L} \propto \frac{\sigma}{R^4} \tag{2.1}$$

where P_t is the transmitted power, G the radar antennae gain, λ the electromagnetic radiation wavelength, σ target radar cross section or RCS, R target range, L radar losses. Most of the parameters in Equation 2.1 are related to the radar sensor itself and can be either measured or estimated, and can be considered constant for a particular radar system operation. These parameters are P_t, G, λ and L. However, σ, the RCS depends on the target (material, size, etc.) and the target's relative orientation with respect to the radar. This means that echos from the same target at the same range but received from different angles might return a different amount of power.

Landmark identification concerns the detection of signals from noisy measurement data. Targets of interest in radar data are usually embedded in noise and clutter. Thus, landmark detection is necessary to identify landmark signals from the noisy power measurement data.

The following sections detail radar data representations used in this thesis work, stochastic detection techniques, and noise reduction methods.

2.2.1 Radar data representation

Returned power received by the radar is usually digitalized for further processing, storage and visualization. Two different sources of radar data were used in this dissertation.

The first source corresponds to a 94 GHz FMCW scanning radar manufactured by Acumine [WCBS+06a], available at the robotics laboratory of the Engineering faculty at Universidad de Chile. This radar was used to capture data from a public park environment in Santiago, Chile. More specifications of this sensor are listed in appendix C.1

The second source is associated to Synthetic Aperture Radar (SAR) data. This data has been produced by NASA/JPL's Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR) mission. The UAVSAR radar is a L-band sensor capable of producing polarimetric and interferometric radar data. The data retrieved corresponds to a location near Sacramento, CA, USA. Appendix C.2 lists some technical details of the system.

Specific details of the different radar data is described in the following sections.

Scanning radar data

The simplest way of presenting radar data is in the form of the received power versus range at a particular bearing angle, also known as an *A-Scope* (Figure 2.1).

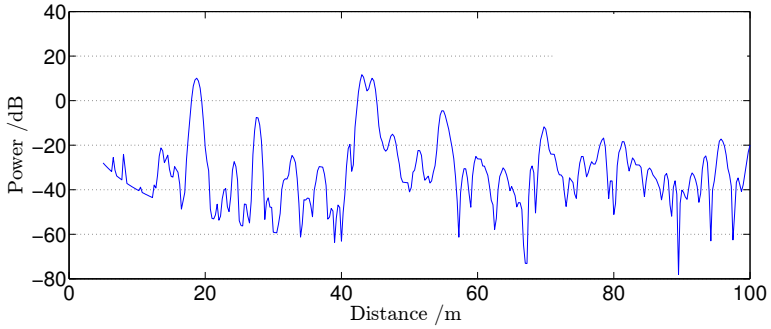

FIGURE 2.1: Received power versus range A-Scope.

If the radar beam is scanned about a vertical axis through 360°, the A-scopes can be combined to form a B-scope (Figure 2.2) which corresponds to an image with range as its vertical axis and bearing angle as the horizontal axis, and the received power value is represented as pixel intensities.

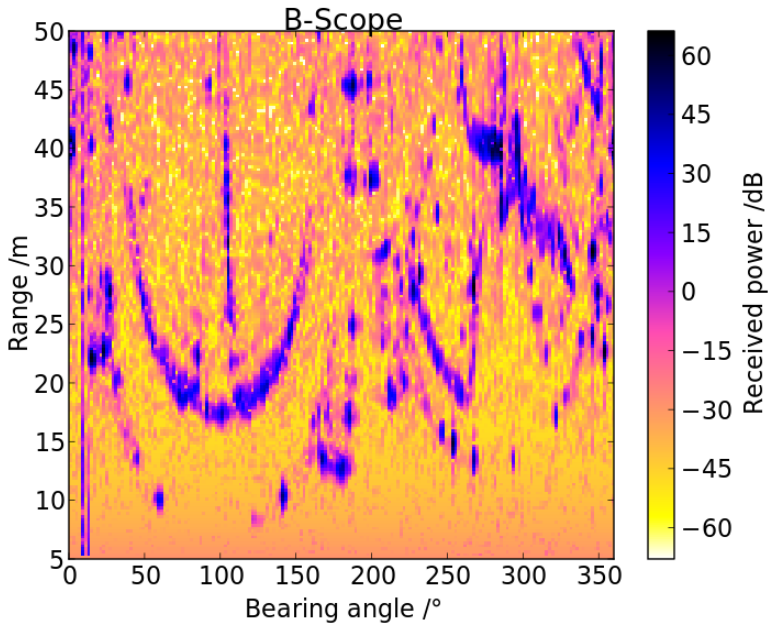

FIGURE 2.2: Received power versus range-bearing B-scope.

The B-scope can be projected into a Cartesian plane using a polar coordinate transformation, known as a *plan position indicator* (PPI) as in Figure 2.3.

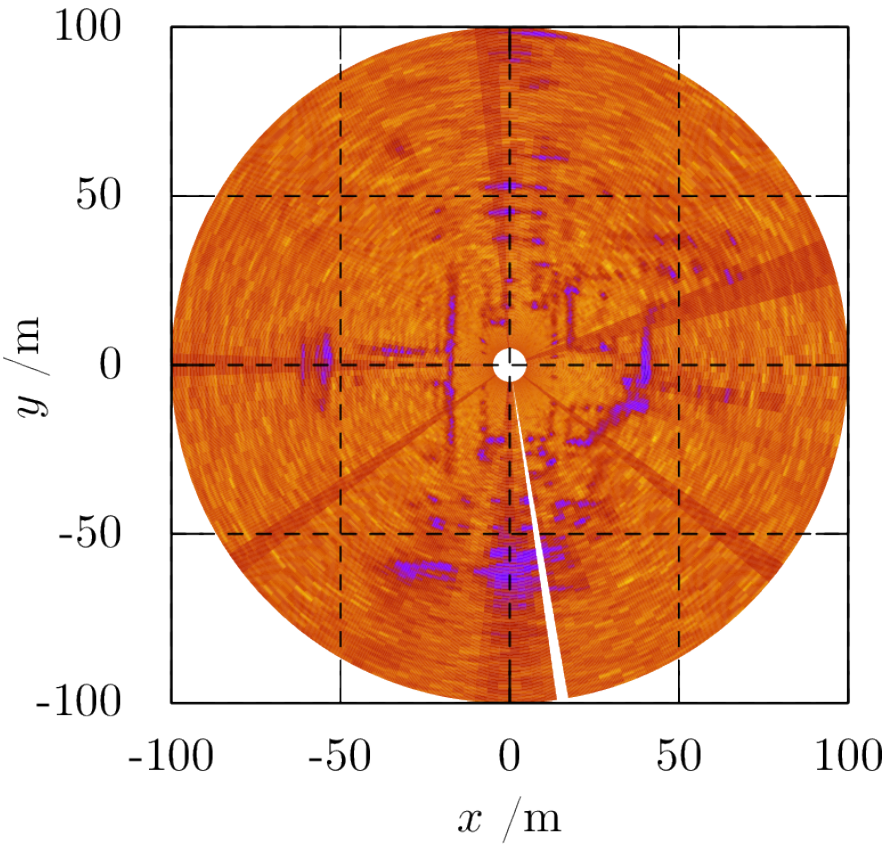

FIGURE 2.3: Plan position indicator (PPI).

SAR data

Synthetic Aperture Radar (SAR) is a side-looking radar[2], which simulates an extremely large antenna or aperture electronically, using the platform flight path. Oftenly, SAR images are thus taken from airborne or spaceborne vehicles. SAR images are mostly used in remote sensing and mapping, usually to generate digital elevation models (DEM) which represent the topography of the scanned surface. Some SAR systems encode radar wave polarization data in the SAR image's color channels. Thus, polarimetric[3] techniques are used to infer some information about the materials reflecting the radar waves. Figure 2.4 shows a polarimetric SAR image taken by the UAVSAR mission over the Gulf of Fonseca in Central America.

[2] *side-looking radar* refers to an aircraft or satellite mounted imaging radar pointing perpendicular to the flight direction.

[3] the measurement and interpretation of transverse (electromagnetic) wave's polarization is called *polarimetry*.

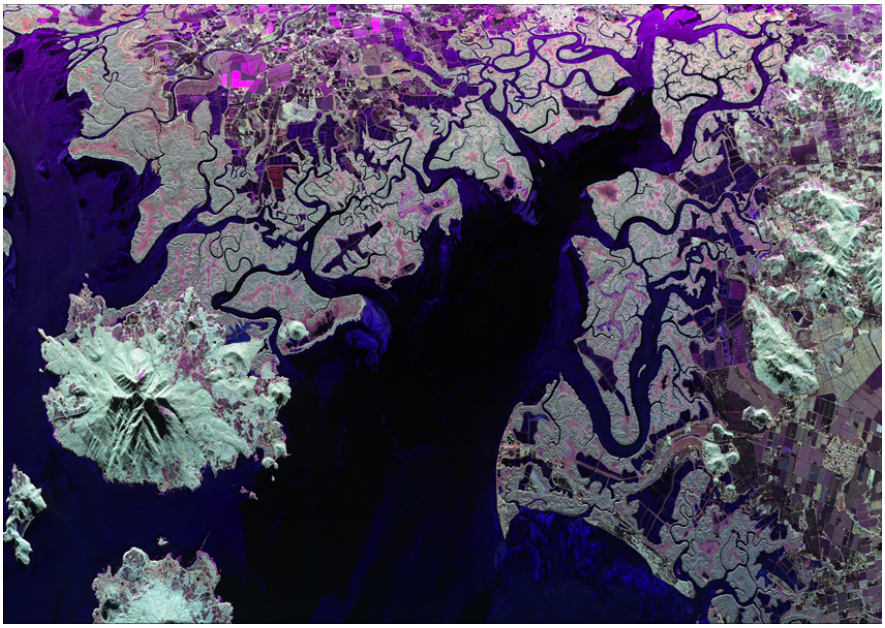

FIGURE 2.4: Polarimetric image of Gulf of Fonseca, Central America (Feb 2, 2010)

2.2.2 Stochastic adaptive detection with CFAR processors

Adaptive, stochastic, landmark detection techniques offer principled methods of detection based on a predefined acceptable probability of false alarm and quantifiable probabilities of detection. The following sections describe some of the most widely used stochastic detection methods used in radar.

CFAR

The Constant False Alarm Rate (CFAR) concept refers to a family of adaptive algorithms widely used in radar to detect target returns against a background of noise, clutter and interference.

The starting point for the processing of radar data is the standard assumption that the target absence and presence distribution types are known, but not their moments, since these typically vary and should be estimated. In this case a test statistic must be derived. If such a statistic can be found, which is independent of any unknowns, then the detection method is known as a CFAR detector.

In most radar signal processing literature to date, a Cell Averaging (CA) CFAR detector is the preferred method of target detection [GK94, FBW01b, Lan96]. Figure 2.5 shows an adaptive CA-CFAR threshold $S_{\log}^{\text{CA-CFAR}}$ (dashed line) applied to the A-scope of Figure 2.1.

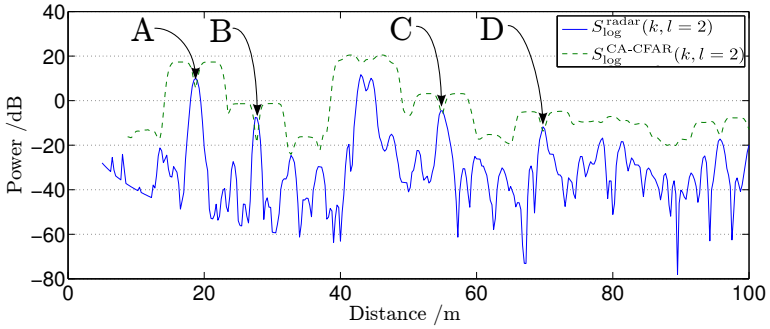

FIGURE 2.5: Sample CFAR threshold superimposed on a power versus range A-scope.

CA-CFAR

The adaptive threshold $S_{\text{lin}}^{\text{CA-CFAR}}$ necessary for CA-CFAR based detection is derived as follows.

[Bar05] shows that the detection probability of a Rayleigh fluctuating target, embedded in exponential noise or clutter, can be determined from the CA-CFAR parameters

$$P_D^{\text{CA-CFAR}}(q) = \left[1 + \frac{\tau^{\text{CA-CFAR}}}{W_f}\left(\frac{1}{1+\widehat{\eta}^{\text{SNP}}(q)}\right)\right]^{-W_f} \tag{2.2}$$

where W_f is the size of the CFAR window and $\tau^{\text{CA-CFAR}}$ is defined as

$$\tau^{\text{CA-CFAR}} = W_f\left(\left(P_{fa}^{\text{CA-CFAR}}\right)^{\frac{-1}{W_f}} - 1\right) \tag{2.3}$$

and $\widehat{\eta}^{\text{SNP}}(q)$ is the estimated received SNP calculated as

$$\widehat{\eta}^{\text{SNP}}(q) = \frac{S_{\text{lin}}^{\text{radar}}(q)}{T(S_{\text{lin}}^{\text{radar}}(q))} \tag{2.4}$$

where $S_{\text{lin}}^{\text{radar}}(q)$ is the linearised received power from the radar in the q^{th} (bearing-range) bin, referred to as "cell under test" (CUT). The adaptive threshold is then defined as

$$S_{\text{lin}}^{\text{CA-CFAR}}(q) = \tau^{\text{CA-CFAR}} \cdot T(S_{\text{lin}}^{\text{radar}}(q)) \tag{2.5}$$

Hence, for chosen values of W_f and $P_{fa}^{\text{CA-CFAR}}$, the $P_D^{\text{CA-CFAR}}$ of a target, within a single received power versus range A-scope, as a function of the estimated received SNP $\widehat{\eta}^{\text{SNP}}$, can be derived. In general, in order to assure that the cells in the window contain noise information only, "guard-cells" are selected around the CUT, accounting for the fact that

some target power can be spread to adjacent cells of the CUT. In this case, another parameter G corresponding to the number of guard-cells is included. By using the following change of variable

$$W_g = W_f - G \tag{2.6}$$

and replacing W_g for W_g for all the CA-CFAR equations, the guard-cells parameter G is incorporated into the method.

In the case of scanning radar data where B-scopes are available, the above (1D) CA-CFAR method could be applied to the range or the bearing angle dimension. Another alternative is to extend it to 2D and process the range and bearing dimensions simultaneously, thus jointly analyzing all available information. Also from a computational point of view it has the advantage of processing the radar data once, instead of sequentially for range, then bearing and integrating the results at the end. The 1D window used to estimate the test statistic is extended to a 2D area surrounding the cell under test (CUT) as shown in Figure 2.6.

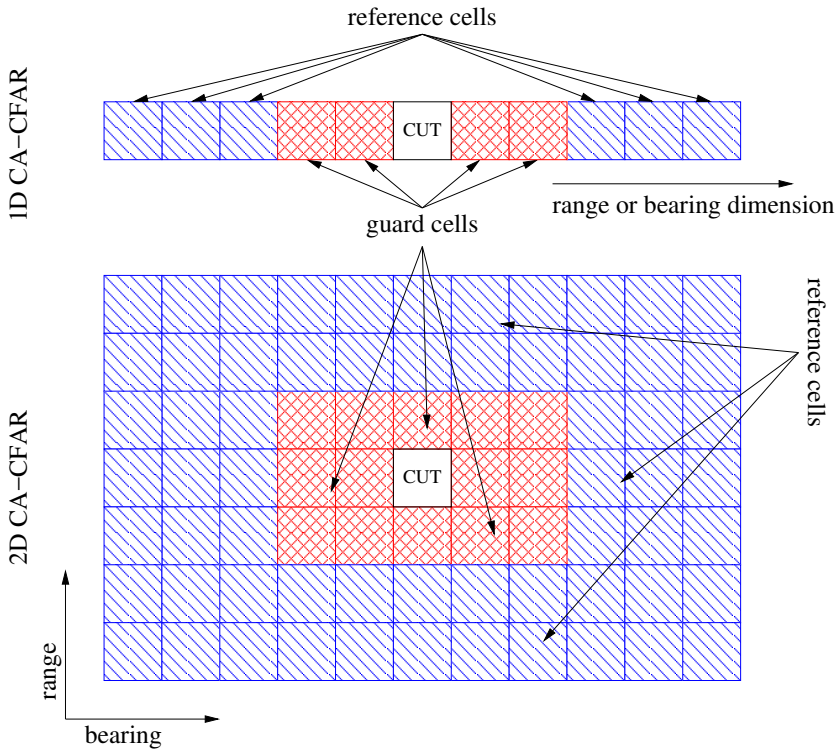

FIGURE 2.6: 1D and 2D CA-CFAR reference windows.

Let W_r and W_b represent the window size in range and bearing, respectively, and G_r and G_b the number of guard-cells also with respect to range and bearing. Then the W_g cells surrounding the CUT in all CA-CFAR equations are replaced by $(W_r+1)\times(W_b+1)-(G_r+1)\times(G_b+1)$ cells surrounding the CUT. By introducing the former changes and $W_G^{2\mathrm{D}}$ defined

as

$$W_G^{\text{2D}} = (W_r + 1)(W_b + 1) - (G_r + 1)(G_b + 1) \tag{2.7}$$

Equation (2.3) and Equation (2.2) now become

$$\tau = W_G^{\text{2D}} \left(\left(P_{fa}^{\text{CA-CFAR}} \right)^{-\frac{1}{W_G^{\text{2D}}}} - 1 \right) \tag{2.8}$$

$$P_D^{\text{CA-CFAR}}(q) = \left[1 + \frac{\tau}{W_G^{\text{2D}}} \left(\frac{1}{1 + \widehat{\eta}^{\text{SNP}}(q)} \right) \right]^{-W_G^{\text{2D}}} \tag{2.9}$$

As can be seen from the equations and their respective plots in Figure 2.7 the 2D CA-CFAR probability of detection is higher than the 1D CA-CFAR detection probability, for a given received signal to noise power ratio η^{SNP}.

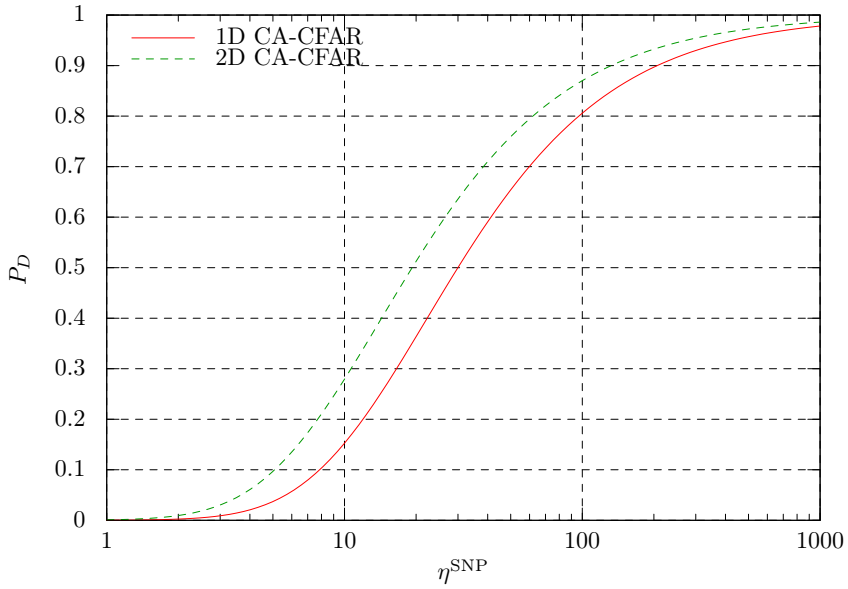

FIGURE 2.7: 1D and 2D CA-CFAR probability of detection. The following CA-CFAR parameters were used: $W_f = W_r = W_b = 20$, $G = G_r = G_b = 4$, $P_{fa}^{\text{CA-CFAR}} = 1 \times 10^{-6}$.

Several other CFAR methods have been developed and current research focuses in CFAR methods with adaptive parameters [GLZ+09]. In particular, the Ordered Statistic (OS) CFAR has been reported to perform well for large targets (with respect to the spatial resolution) and in SAR images, due to their noise and clutter being usually modeled by Weibull or K distributions, and the higher effect of multiplicative speckle noise present in such images.

OS-CFAR

In the OS-CFAR method the test statistic $T(S_{\text{lin}}^{\text{radar}}(q))$ is obtained by choosing the k^{th} value from the ordered set of power values in the CFAR window

$$S_{\text{lin}}^{(1)} \leq S_{\text{lin}}^{(2)} \leq \ldots \leq S_{\text{lin}}^{(k)} \leq \ldots \leq S_{\text{lin}}^{(W_f-1)} \leq S_{\text{lin}}^{(W_f)} \tag{2.10}$$

a value of $k = \frac{3W_f}{4}$ has been suggested in [Roh83] to represent a good estimate for typical radar applications. The parameter $\tau^{\text{OS-CFAR}}$ needs to be calculated numerically from

$$P_{fa}^{\text{OS-CFAR}} = \prod_{i=0}^{k-1} \frac{W_f - i}{W_f - i + \tau^{\text{OS-CFAR}}} \tag{2.11}$$

while the probability of detection is obtained from

$$P_D^{\text{OS-CFAR}} = \prod_{i=0}^{k-1} \frac{W_f - i}{W_f - i + \frac{\tau^{\text{OS-CFAR}}}{1+\widehat{\eta}^{\text{SNP}}(q)}} \tag{2.12}$$

Unfortunately, the noise and target distribution assumptions in CFAR are often violated, yielding higher false alarm and missed detection rates than those theoretically derived. To reduce this problem, and if time is available to acquire multiple scans of a fixed environment, it makes sense to exploit the high target correlation between scans to further reduce the uncertainty in the existence of objects and reduce the noise. Techniques which implement this concept are generally referred to as integration methods.

Binary Integration

A simple but effective integration method widely used in the radar community is the Binary Integration (BI) Method [Bar05].

For the integration of L cells corresponding to a particular range and bearing obtained in successive scans, the number of power values in cells which exceed the single observation threshold (i.e. CA-CFAR threshold) is counted. If this number is greater than or equal to a chosen value M^{BI}, a detection is declared. This detector is then called an M^{BI} out of L binary integrator. A more detailed description now follows.

Given L and a desired resulting probability of false alarm P_{fa}^{BI}, an optimum or near optimum M^{BI} has to be determined. Using an indicator function $I(q,l) = 1$ when a single scan detection is declared and $I(q,l) = 0$ otherwise, the M^{BI} out of L binary integrator

declares a detection when:

$$\sum_{l=1}^{L} I(q, l-1) \geq M^{\,\mathrm{BI}} \tag{2.13}$$

q corresponds to an index identifying a particular range and bearing, and l corresponds to the scan number. When integrating L scans, the probability of detection [Bru90] yielded by this method is

$$P_D^{\mathrm{BI}} = \sum_{l=M^{\,\mathrm{BI}}}^{L} \frac{L!}{l!(L-l)!} \Big(P_D^{\text{CA-CFAR}}\Big)^{l} \Big(1 - P_D^{\text{CA-CFAR}}\Big)^{L-l} \tag{2.14}$$

where $P_D^{\text{CA-CFAR}}$ is the probability of detection in a *single scan*, $1 - P_D^{\text{CA-CFAR}}$ being the probability of a miss-detection, and $M^{\,\mathrm{BI}} < L$ is the optimal BI parameter for a given L. Likewise, if $P_{fa}^{\text{CA-CFAR}}$ is the probability of false alarm in a single scan, then the probability of false alarm for the binary integration method is

$$P_{fa}^{\mathrm{BI}} = \sum_{l=M^{\,\mathrm{BI}}}^{L} \frac{L!}{l!(L-l)!} \big(P_{fa}^{\text{CA-CFAR}}\big)^{l} \big(1 - P_{fa}^{\text{CA-CFAR}}\big)^{L-l} \tag{2.15}$$

for $1 \leq M^{\,\mathrm{BI}} \leq L$, and $0 \leq P_{fa}^{\text{CA-CFAR}} \leq 1$

From Equation (2.15) it can be seen that for a desired false alarm rate P_{fa}^{BI}, the false alarm requirement for the single-scan CA-CFAR detector is relaxed, i.e. $P_{fa}^{\text{CA-CFAR}}$ for a single component scan can now be greater than P_{fa}^{BI}, so that the $P_D^{\text{CA-CFAR}}$ is also increased. This in turn, yields a P_D^{BI} higher than the single observation $P_D^{\text{CA-CFAR}}$. Figure 2.8 shows detection probability curves for the single observation CA-CFAR and different $M_{\text{opt}}^{\mathrm{BI}}$-of-$L$ integrators. Over multiple scans, the probability of detection is increased with the number of observations available.

BI offers a robust technique to identify which parts of signal are noise. The detection probability obtained by means of the BI method will be used as a target existence probability estimate in the radar noise reduction method presented in section 3.

2.2.3 Noise reduction techniques

When time is available to obtain multiple sensor data readings from a static scene at the same location, it is possible to exploit the correlation in the sequence of data readings to reduce noise, and consequently improve detection. B-Scope or PPI radar data is usually represented by images which make noise reduction techniques suitable to be used in these kinds of radar data. Several methods have been developed in the field of image processing to

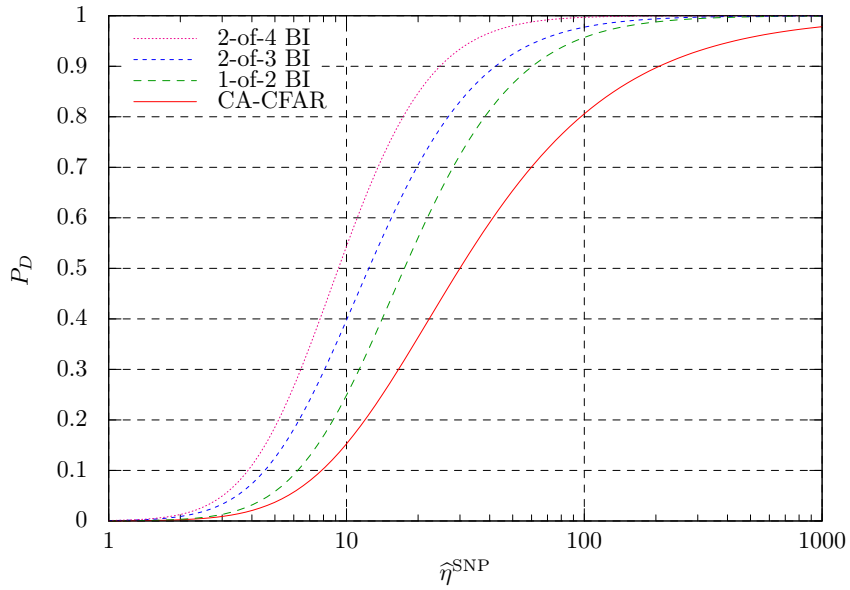

FIGURE 2.8: BI and CA-CFAR probabilities of detection. The 1D CA-CFAR parameters $W = 20$, $G = 4$ and $P_{fa}^{\text{CA-CFAR}} = 1 \times 10^{-6}$ were used.

reduce noise in both stationary and dynamic image sequences in applications as diverse as object tracking surveillance, autonomous navigation, motion analysis, and astronomical and medical imaging [BKE+95]. A sequence of 2D images is represented by a 3D volume where the third dimension corresponds to the temporal dimension or the sequence index. Many of the methods to process such 3D signals have been developed by generalizing well known 2D filtering techniques by extending the support of a filter in the temporal domain. The classical Wiener filter [Wie64], extended to a 3D form, is an example of such an approach. The Wiener filter is a linear time-invariant estimator which adopts a *Minimum mean square error* (MSE) statistical approach. Adaptive noise canceling, developed as a variation of the original Wiener optimal filtering theory, was presented in [WGM+75]. The adaptive noise cancel-ling application uses a reference signal correlated with the noise to obtain a noise estimate. This estimate is then subtracted from the noisy signal. An application of Wiener filtering to 3D medical imaging data [MFALRAW07] extends the classical Wiener implementation by estimating the filter parameters using a sequence of observations based on the calculation of local statistics (calculated in a small window around the point of interest). In radar applications, a 2D Wiener filter, also using local statistics, has been used to reduce noise in weather radar data [DT12].

Work by Donoho and Johnstone [Don95, DJ94b, DJ94c] introduced the denoising capabilities of the Wavelet transform. The basic method is *Denoising by Thresholding* [BGG98]. It is analogous to frequency domain filtering based on the Fourier transform. The wavelet

time-frequency approach however, attempts to reduce noise by preserving a number of coefficients associated with components with high information energy, and discarding the rest. It is assumed that noise (often considered to be additive Gaussian) is spread homogeneously among all signal frequency components. Thus by discarding the coefficients of the components not highly correlated with the signal, a significant amount of noise is eliminated. Coefficients with a magnitude higher than the threshold are considered to hold mostly signal information, and those lower than the threshold are considered to carry mainly noise energy. In radar applications several articles have demonstrated the use of wavelets to reduce noise. Chen [CC96] proposed a recursive thresholding method for radar image denoising, while Aly demonstrated the use of wavelet packet transforms and higher order statistics to detect and localize RF radar pulses in noisy environments [AOE06]. In general, most noise reduction algorithms (both 'classical' and wavelet-based) assume the noise to be additive Gaussian [BKE+95, BGG98], which is useful for a broad range of applications. However, in radar imaging, the Gaussian noise assumption is not realistic. Another critical aspect of wavelet denoising is the appropriate threshold selection. An adaptive threshold method was introduced by Chen [CH05], which adapts the threshold to the coefficients' statistics, relaxing the Gaussian assumption of most wavelet based methods. Another adaptive method, presented by Jin [JFWJ03] uses abrupt changes in the signal to adapt weights to calculate local means and variances. This approach is reported to reduce the ripple-like artifacts usually found around edges when using wavelet denoising techniques.

Spectral noise subtraction [Bol79] (or noise subtraction) is another well known technique used in signal processing to reduce noise. This method requires an estimate of the noise power spectrum in order to perform the noise subtraction from the original noisy data.

The following sections describe these three noise reduction methods.

Wiener filter

This method of noise reduction corresponds to the application of a discrete-time minimum-mean-square-error filter. Such a filter is known as a Wiener-Kolmogorov filter or Wiener filter for short [Wie49, Kol78]. Wiener filter noise reduction is equivalent [Ein12] to estimating a signal $Y(z)$ corrupted by noise $V(z)$ from the measurement $U(z) = Y(z)+V(z)$. Let $H_E(z)$ be the (Wiener) filter which minimizes the error $E(z) = ||Y(z) - \widehat{Y}(z)||^2$, where $\widehat{Y}(z)$ represents an estimate of the original signal $Y(z)$, and z is the discrete frequency-domain variable (see Figure 2.9).

Wiener filters assume additive noise and that the signals are stationary, linear stochastic processes. Because radar images, as well as natural images, consist of smooth areas, textures

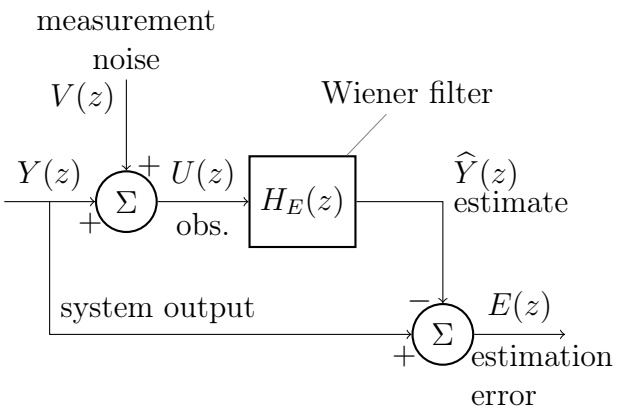

FIGURE 2.9: Estimation problem and the Wiener filter.

and edges, they are not *globally* stationary, but can be treated as *locally* stationary. This led to the derivation of the Lee filter [Lee80], which has been extensively used in video denoising, where it has proved to be successful in terms of noise removal and preserving some important image features (e.g. edges) [JFWJ03]. The Lee filter assumes that all samples within a local window are from the same structure (local stationarity). This assumption is invalid when sharp edges are encountered within the window, therefore the mean is blurred and the variance increases near the edges, which results in a degraded image near the edges.

Wavelets

Some of the limitations of the Wiener filter, particularly its inability of preserving non-linear features in the data, can be overcome by using non-linear filters. However, in general, finding the parameters for a non-linear filter is a complex task. Since the introduction of the wavelet denoising methods by Donoho and Johnstone [DJ94b, DJ94c, Don95] a powerful, yet simple to implement non-linear filter for noise reduction has become available and is widely used in practical implementations.

Given the discrete wavelet transform (DWT) coefficients, which are analogous to the Fourier transform coefficients but in a 2D (time-frequency) space, the main idea behind the wavelet denoising scheme is that most of the signal's energy is concentrated in a few coefficients, while the noise energy is spread across all coefficients[4]. This effect is known as time-frequency localization. Thus, by preserving the coefficients which hold most of the signal's energy and setting the remaining ones to zero, the signal reconstructed through the inverse DWT using these new coefficients will be a close approximation of the original signal with reduced noise.

[4]The way the noise is distributed across the coefficients is highly dependant on the type and distribution of the noise.

The application of the DWT and its inverse can be interpreted as a signal being processed through a filter bank, as shown in Figure 2.10.

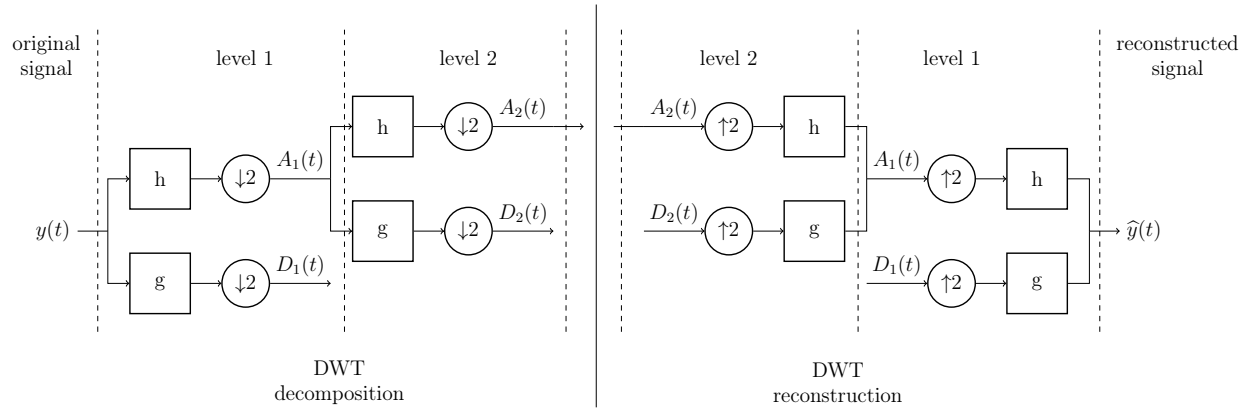

FIGURE 2.10: Wavelet decomposition and reconstruction as a filter bank.

Wavelet denoising estimators are derived under the *minimax* criterion. This corresponds to finding an estimator whose maximal risk is minimal among all estimators, as opposed to the minimum MSE criterion used in the Wiener filter. In the wavelet denoising approach, noise and signal spectra might overlap in the frequency domain, but the coefficients associated with the signal and those related to noise are expected to have different amplitudes. Therefore clipping, thresholding or shrinking of the amplitude of the transform can be used to separate signals or remove noise. A thresholding function corresponding to the estimated noise level, which preserves signal values above the threshold has to be chosen.

The particular thresholding function used in this work is the universal threshold proposed by Donoho and Johnstone [DJ94a, LZ12]

$$T = \sigma\sqrt{2\log_e N};\ \text{(where } N \text{ the signal size and } \sigma^2 \text{ the noise variance)} \tag{2.16}$$

with the soft-thresholding method (refer to Eq. (2.17) and Figure 2.11) proposed in [Don95].

$$y = \begin{cases} x - T & \text{if } x \geq T \\ x + T & \text{if } x \leq -T \\ 0 & \text{if } |x| < T \end{cases} \tag{2.17}$$

This method generates an estimated signal with a smaller amplitude than the original one, but it retains the regularity[5] of the signal.

[5]regularity corresponds to areas in which the signal is continuous, while discontinuities correspond to irregular points

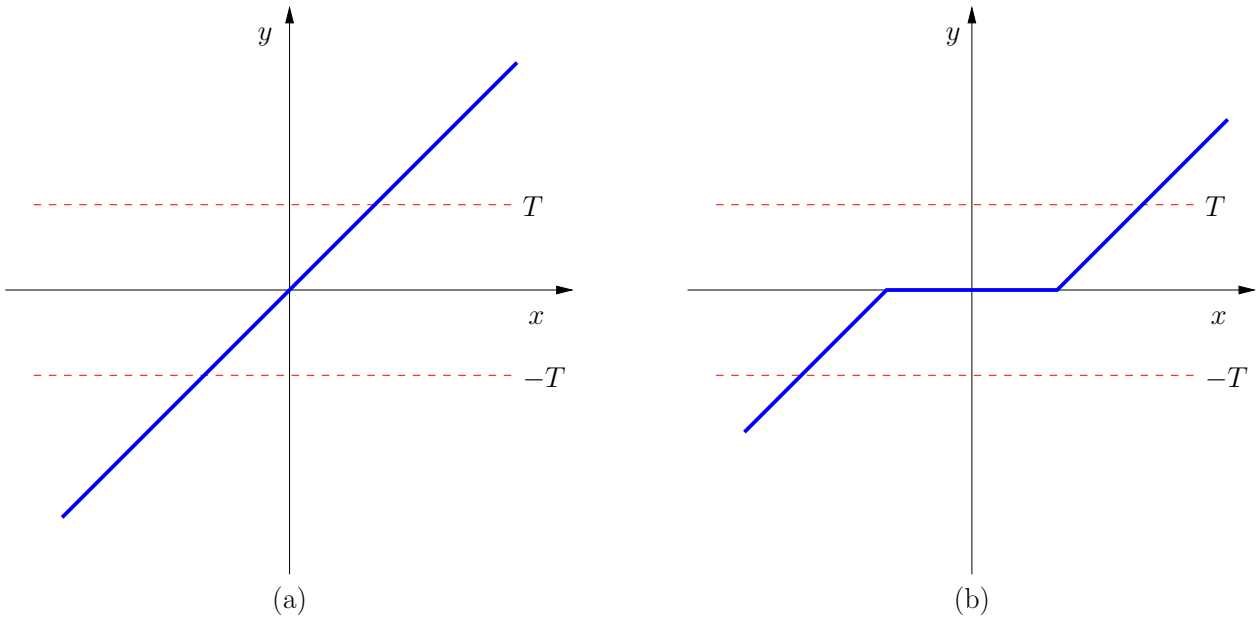

FIGURE 2.11: Soft thresholding: (a) original values, (b) soft thresholding applied to values.

Spectral noise subtraction

The spectral noise subtraction technique, first presented by Boll [Bol79], is based on estimating the magnitude frequency spectrum of the underlying signal of interest by subtracting the noise magnitude spectrum from a noisy signal. An estimate of the noise spectrum is required. Such an estimate is approximated by averaging the noise power using parts of the noisy signal which are known to hold no information of interest. Further modifications to enhance the noise subtraction method were presented by Berouti et al. [BSM79] which reduce broadband and background noise.

2.3 Feature detection in 3D environments

Applications such as robot localization, simultaneous localization and mapping (SLAM), augmented reality and 3D surveying involve the use of sensors to capture the environment. Three-dimensional data can be generated using the output of different sensing technologies, such as stereo cameras, structured or modulated light sensors, scanning range finders, etc. Laser based scanning range finders capable of generating 3D ladar data are widely used in robotic mapping and surveying applications, due to its relatively simple approach to obtain the 3D coordinates, in contrast to stereo images calculations, for instance. Also, they are preferred for its range of operation and high measurement precision. Ladar sensors output a

set of 3D points (in Cartesian, spherical, cylindrical or other 3D coordinate system, depending on the sensor). The points correspond to objects in the environment which are illuminated by the laser beam from the laser range finder and reflect a sufficiently large amount of laser energy back to the sensor. Some ladar sensors are able to attach to each point extra information such as the reflected laser intensity or the corresponding color in the light spectrum (as a Red, Green, Blue triplet) captured with an optical camera. This set of points is referred to as a 3D point cloud. Although other sensors are capable of generating 3D point clouds, without loss of generality this section will focus on ladar data.

Because most objects of interest reflect the majority of the laser energy back to the sensor, a ladar sensor can only detect those objects in direct line-of-sight. To be able to detect the shaded parts of the environment, the sensor has to scan it from different perspectives, each generating a new 3D point cloud. Therefore, it is usually required to match successive scans from the environment, often from different relative locations. In some applications, this process is referred to as registration. In order to achieve a correct matching, many methods rely on features present on the environment, which can be identified in the different scans, and then used as common references. These features need to be detected or calculated from the sensor data. One approach is to use particular properties which naturally occur in the environment, thus they are called natural features. Often, natural features are not available or their quality is not enough to yield a correct matching. In this situations artificial markers are incorporated to the environment. The state of the art in both approaches is described next.

2.3.1 Natural features

Fully automatic algorithms favour the automatic extraction of natural features from scans, which are then exploited to find the correspondences of scan pairs. Lee et al. [LKC04] presented a matching method in order to find the correspondences of features in two omni-directional images. Hansen et al. [HCBD07b, HCBD07a] considered an application of scale invariant feature detection using scale-space analysis suitable for use with wide field of view cameras.

The reflectance intensity provided by modern laser scanners gives additional information which can be used for registration. Based on this concept, Böhm and Becker [BB07] suggest the use of SIFT [Low04] features for automatic registration. More recently, Endres et al. [EHS+14] compare SIFT, SURF [BETVG08] and ORB [RRKB11] as feature descriptors for visual SLAM with a RGB-D camera. Weinmann et al. [WWHJ11] and Weinmann and Jutzi [WJ11] present a method that is based on both reflectance and range images. All these

approaches find the correspondences for registration based on the features extracted from the images and directly from the point cloud.

Other approaches rely on 3D structure. Brenner et al. [BDR08] use 3D planar patches and the normal distribution transform (NDT) on several 2D scan slices for registration. These approaches use the global features of the scans, i.e. features that describe the complete scan.

Several researchers use features that describe small regions of the scan for place recognition and the registration of scans (Huber [Hub02], Steder et al. [SGB10], Barnea and Filin [BF08]).

Houshiar et al. [HEBN13] present an approach to convert 3D scans into panoramic images, extract features from these images, use feature descriptors for data association and compute the transform using RANSAC and a least squares method.

2.3.2 Artificial markers

Marker based registration uses artificially defined landmarks as corresponding points. This form of supervised data association ensures that the scans are registered at their correct location. These algorithms can extract the 3D position of the markers either manually or automatically.

Nevertheless, although the techniques to extract natural features continue to improve, there are situations in which artificial markers are required or show a superior performance [Fia10]. These include the cases in which the environment is highly dynamic and natural features are not consistent between successive scans [SKC13, NHY$^+$00], when natural features are not present in the sufficient quantity and uniqueness [Fia10] or a high accuracy and reliability is required [Fia10, NHY$^+$00]. Moreover, natural features often require large computational times to extract such features. Applications where artificial landmarks are used include indoor augmented reality, pose determination in industrial settings, reverse engineering of 3D objects, real-time tracking systems, face recognition, medical imaging, sensor calibration and 3D laser surveying [Fia10, BPR12, BGG$^+$04].

Due to the lack of adequate analytic approaches in literature, markers are specifically designed for each task. A proper marker design aims at a highly reliable detection and and a very low probability of false alarms [Fia10].

Markers designed for the augmented reality (AR) application are usually detected by video camera. Most of them consist of high contrast (black and white) patters, encoding some information, such as an identification tag, instructions or commands to be processed after detection, orientation marks, etc. The standard method to detect these markers begin with a binarization step based on thresholding, followed by connected component analysis to locate blobs where a unique feature can be detected by means of shape or line detection. The

ARTag marker presented in [Fia10] replaces the simple constant threshold method, widely used in AR markers, with a gradient based edge detection to avoid the dependency on lighting variations and to add the ability to detect partially occluded markers using heuristics. After detection, AR markers are subject to a second stage for verification and identification of the information embedded in the marker.

Artificial markers are also used align 3D point clouds for reverse engineering of objects captured with a 3D scanner such as a stereo vision system as reported in [BPR12]. Natural features, usually shape descriptors from the scanned object surface are not available due to the properties of the surface (flat or uniform curvature geometries). The markers used in the presented method are patterns of black and white stripes or circular black and white markers with similar properties as those used for the AR application mentioned above.

In robot localization and robot SLAM the use of artificial landmarks is required to achieve high accuracy in highly dynamic environments or when there are few natural features available. When laser sensors are used on board a robot or automated vehicle the artificial landmark used is usually a laser reflecting surface [ROSF11]. A global localization method for automated guided vehicles (AGV) in an industrial environment using reflective markers is presented in [ROSF11]. The method is able to work in large industrial environments with several hundred markers, which surpass the capability of commercial off-the-shelf system handling only a few dozen markers at a time. The method assumes the location of the markers in the map and the map itself are known *a priori*. In an off-line phase run once at the beginning to calibrate the position of the reflectors and thus, reduce the effect of false alarms (outliers) from reflecting surfaces other than the actual landmarks. These outliers are removed using an heuristic which converges with fewer iterations than comparable RANSAC-based methods.

Another example of SLAM using artificial retro-reflective markers is presented in [NHY$^+$00]. A robot is used to monitor children activities in a Kindergarten environment. The children also interact with the robot, thus a high accuracy SLAM is required to avoid accidents. The reported SLAM approach achieve an average error of 24 mm, corresponding to an error rate of 0.14% of the total running distance. The markers' design consists of of 5 stripes of alternating reflective and non-reflective material. The detection method finds wall lines in the environment using a Hough transform, and search for the reflective pattern [NS12] on the detected lines.

In 3D laser surveying application it is required to rapidly and automatically detect points marked by retro reflectors. The process and device described in US Patent 6,734,952 [BGG$^+$04] assigned to Leica Geosystems is used to detect retro reflective markers using a geodetic measuring system (total station, theodolite or similar device). The device

includes a specific hardware called automatic locating unit (AZE[6]). The AZE unit is responsible of marker seeking, positioning and marker point measuring. Once the AZE has detected the markers, their positions are stored along the actual data measured by the device. There are multiple approaches to build an AZE. Most AZE units are not robust with respect to the detection of foreign objects reflecting sunlight or the AZE's laser beam. The unit disclosed uses two fan-shaped laser beam of narrow spectrum to scan on wide field of view for reflective surfaces. The narrow spectrum allows the AZE to reject false reflections of sunlight. The rejection of foreign objects reflecting the sensor's laser is done by specific heuristics depending on the estimated distance of the target. For close targets (less than 10 m) a special arrangement of the fan axes allows to discard false detections. When the distance to the target is larger than 10 m a two dimensional intensity image is used. The target's widths are calculated from this image and thus validated as a marker if they fall within tolerance values. This method requires a considerable amount of time during the setup phase to detect the markers before starting the normal 3d surveying procedure.

2.4 Optimal Sub-pattern Assignment (OSPA) map metric

Target detection is a state estimation problem. As such, the concept of estimation error is of great importance. Many methods used to compare estimated detections with ground truth information are only defined if the estimated set has the same number of objects as the ground truth set [BAL+17], such as root mean squared (RMS) and normalized estimation error squared (NEES) metrics [BSLK04, HHD10, CMCTN07, WEL07]. Nevertheless, in most real applications the number of detected objects and the number of ground truth objects will be different, which often lead to pruning the estimated state or the ground truth vectors commonly used to represent these two sets. The metric used here (which jointly considers errors in target location, and number estimates) is based on a p^{th} order Wasserstein construction [SVV08]. This metric has been recently introduced and has been shown to produce more consistent results than others (Hausdorff, OMAT) in multi-object systems. It is formulated in terms of the ground truth set of targets $\widehat{\mathcal{M}}$ and the detected set $\mathcal{M}$. If $|\mathcal{M}| > |\widehat{\mathcal{M}}|$, it is given by,

$$\bar{d}^{(v)}(\widehat{\mathcal{M}}, \mathcal{M}) = \left(\frac{1}{|\mathcal{M}|} \left(\min_{\pi \in \Pi_k} \sum_{i=1}^{|\widehat{\mathcal{M}}|} d^{(v)}(\widehat{m}^i, m^{\pi(i)})^p + v^p(|\mathcal{M}| - |\widehat{\mathcal{M}}|) \right) \right)^{1/p} \tag{2.18}$$

[6]AZE: from german *Automatische Zielauffindungseinheit*

where π correspond to the permutation in Π_k (the set of permutations on $\{1, 2, \ldots, k\}$) which minimizes the sum of the distances between the elements from $\widehat{\mathcal{M}}$ and $\mathcal{M}$, while

$$d^{(v)}(\widehat{m}^i, m^{\pi(i)}) = \min(v, ||\widehat{m}^i - m^{\pi(i)}||) \tag{2.19}$$

is the minimum of the cut-off parameter, v, and the Euclidean distance between the ground truth target location, $\widehat{m}^i$ and the detected target location $m^{\pi(i)}$. If $|\mathcal{M}| < |\widehat{\mathcal{M}}|$ the metric is obtained through $\bar{d}^{(v)}(\mathcal{M}, \widehat{\mathcal{M}})$. To find the optimal $\pi \in \Pi_k$ is a special case of integer linear programming, which can be solved efficiently using the Hungarian method [Fra04] for optimal point assignment.

Chapter 3

Binary Integration Noise Reduction

In this work a new approach for radar image denoising is introduced and compared to the classical Wiener filter and to the more recent wavelet based denoising approaches. The proposed method's implementation presents a lower computational complexity than both the 3D Wiener filter and the 3D wavelet approaches. The method further reduces the mean noise value when compared with the other two methods. It uses statistical information provided by the Binary integration detector to identify parts of the received signal corresponding to noise. It uses those parts to obtain an estimate of the noise power spectrum by recursive averaging. This noise estimate is then used for power spectral subtraction [Bol79] (or noise subtraction) to reduce noise. In particular, Binary Integration (BI) combines the output of several single-observation detectors to improve the detection probability, while maintaining the desired, acceptable false alarm rate. The single-observation detector used is a member of the Constant False Alarm Rate (CFAR) family of stochastic detectors, widely used in radar [Sko08]. In the case of radar data, the binary integrator's probability of detection can be used to identify sections which have low probability of having any target information and therefore they can be used to estimate the noise magnitude. This new method has been called *Binary Integration Noise Reduction* or BINR for short.

3.1 Integrating BI with spectral noise subtraction

The noise power estimate $\widehat{\Sigma}_n(l)$ can be calculated as in Equation (3.1), which corresponds to a recursive smoother using a fixed parameter α_d, and the binary integration probability of

detection from Equation (2.14).

$$\begin{aligned}\widehat{\Sigma}_n(l) = &\Big(\alpha_d \widehat{\Sigma}_n(l-1) + (1-\alpha_d) S_{\text{lin}}^{\text{radar}}(l)\Big)\big(1 - P_D^{\text{BI}}(l)\big) \\ &+ \widehat{\Sigma}_n(l-1) \times P_D^{\text{BI}}(l)\end{aligned} \tag{3.1}$$

The first term on the right of Equation (3.1) represents the smoothed (averaged) noise power, weighted by $\big(1 - P_D^{\text{BI}}(l)\big)$ during *target absence* sections of the signal, while the second one shows that the previous estimate is preserved and not updated if there is a high probability of *target presence* $\big(P_D^{\text{BI}}(l) \to 1\big)$. Introducing

$$\tilde{\alpha}_d(l) = \alpha_d + (1-\alpha_d) P_D^{\text{BI}}(l) \tag{3.2}$$

Equation (3.1) can be rewritten as

$$\widehat{\Sigma}_n(l) = \tilde{\alpha}_d(l)\widehat{\Sigma}_n(l-1) + \big[1 - \tilde{\alpha}_d(l)\big] S_{\text{lin}}^{\text{radar}}(l) \tag{3.3}$$

In Equation (3.2), $\tilde{\alpha}_d(l)$ is a time varying smoothing parameter. Hence the noise spectrum can be estimated using past spectral received power values, together with a smoothing parameter which itself varies according to the BI probability of detection $P_D^{\text{BI}}(l)$. The adaptive noise estimator is summarized in the block diagrams of Figures 3.1 and 3.2.

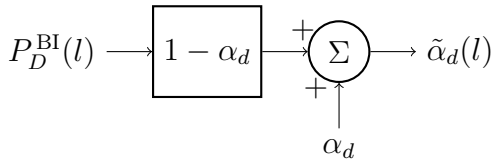

FIGURE 3.1: Block diagram summarising the time varying parameter $\tilde{\alpha}_d$.

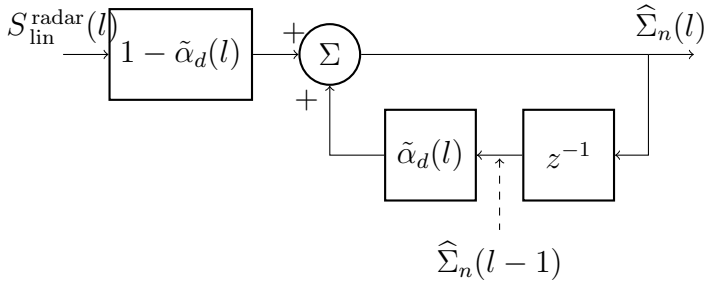

FIGURE 3.2: Binary integration noise reduction block diagram.

The noise power estimate $\widehat{\Sigma}_n(l)$ can be used to obtain a *Binary Integration Noise Reduction* (BINR) power estimate $\hat{S}_{\text{lin}}^{\text{BINR}}(l)$ using the method of power spectral subtraction [Bol79].

In the basic spectral subtraction algorithm, the average noise power, $\widehat{\Sigma}_n(l)$ is subtracted from the noisy range spectrum. In [BSM79], a method which further reduces background noise for SNPs between -5 and 20 dB was devised based on subtracting an over-estimate of the noise power and preventing the resultant spectral components from reaching below a preset minimum level, termed the "spectral floor". This method leads to a reduced noise power estimate $\hat{S}_{\text{lin}}^{\text{BINR}}(l)$ given by

$$\widehat{S}_{\text{lin}}^{\text{BINR}}(l) = \begin{cases} S_{\text{lin}}^{\text{radar}}(l) - c\widehat{\Sigma}_n(l) & \text{if } S_{\text{lin}}^{\text{radar}}(l) > c\widehat{\Sigma}_n(l) \\ d \times \widehat{\Sigma}_n(l) & \text{otherwise} \end{cases} \tag{3.4}$$

where c is an over-subtraction factor ($c \geq 1$) and d is the spectral floor parameter ($0 < d < 1$). A value of c which is larger than 1 represents the fact that it is necessary to subtract more than the expected value of the noise (which is a smoothed estimate) to make sure that most of the noise peaks are removed. The spectral floor parameter d, when greater than zero, ensures that the remnants of the noise peaks are masked by neighbouring spectral components of comparable magnitude. This results in a reduction of broadband noise, when compared with the original power spectral subtraction method. [BSM79] further demonstrated the advantages of an adaptive over-subtraction factor c, which varies between frames of recorded spectra, as a function of the estimated signal.

3.2 Implementation and computational complexity

This section discusses the implementation and computational complexity, as a function of the data size, of all three noise reduction methods. All the algorithms were implemented using the SciPy signal processing toolbox [JOP$^+$]. The same received radar power signal S is used as the input for all the methods. The signal includes additive noise. It is a 3D array formed by stacking L radar scans. Each radar scan $S(l)$ corresponds to a B-scope of size $N_p = N_r \times N_b$ cells (pixels), where N_r represents the number of range bins (rows) while N_b the number of bearing bins (columns). The first dimension in the 3D array represents range (r), the second represents bearing (b) and the third represents the scan number (l). The sequence of L power values for a particular range and bearing is denoted as $S_{r,b}$ while an individual voxel is indicated by $S_{r,b,l}$. The total number of voxels in S is $N_v = N_b \times N_r \times L$. The complexity analysis will be carried out in terms of N_p and L ($N_v = N_p \times L$), to separate the effects of the radar image size from those related to the number of observations.

3.2.1 Wiener filter implementation and complexity

The Wiener filter noise reduction technique used in this work requires a support volume, in order to calculate the local means and variances. For a given voxel, the cells contained in the support volume around it will be used to calculate the local statistics (see lines 4 and 7 in Algorithm 1) . Let $(2K_r+1), (2K_b+1), (2K_l+1)$ be the dimensions of this volume in the range (rows), bearing (columns) and observation (depth) dimensions. K_r and K_b are related to the expected range and bearing size of targets. If these parameters are set too high the local mean and variance of a target voxel will incorporate not only target values but also undesired noise values. On the other hand, if they are to small, the local mean and variance estimates would be degraded. $2K_l + 1 = L$ has been chosen for the window to incorporate information from all observations at each value in the temporal (l) dimension. The **for** loop

Algorithm 1 Wiener filter

1: **procedure** WIENER(signal S, 3D support volume dimensions K_r, K_b, K_l)
2: $K_v \leftarrow K_r \times K_b \times K_l$
3: **for all** $S_{r,b,l} \in S$ **do** ▷ Estimate local mean
4: $\mu_{r,b,l} \leftarrow \frac{1}{K_v} \sum_{k_r=r-K_r}^{r+K_r} \sum_{k_b=b-K_b}^{b+K_b} \sum_{k_l=l-K_l}^{l+K_l} S_{k_r,k_b,k_l}$
5: **end for**
6: **for all** $S_{r,b,l} \in S$ **do** ▷ Estimate local variance
7: $\sigma^2_{r,b,l} \leftarrow \frac{1}{K_v} \sum_{k_r=r-K_r}^{r+K_r} \sum_{k_b=b-K_b}^{b+K_b} \sum_{k_l=l-K_l}^{l+K_l} \left(S_{k_r,k_b,k_l} - \mu_{r,b,l}\right)^2$
8: **end for**
9: $n_p \leftarrow \text{mean}(\sigma^2)$ ▷ Estimate noise power
10: **for all** $S_{r,b,l} \in S$ **do** ▷ Calculate output
11: **if** $\sigma^2_{r,b,l} < n_p$ **then**
12: $S^{\text{wiener}}_{r,b,l} \leftarrow \sigma^2_{r,b,l}$
13: **else**
14: $S^{\text{wiener}}_{r,b,l} \leftarrow \left(S_{r,b,l} - \mu_{r,b,l}\right)\left(1 - \frac{n_p}{\sigma^2_{r,b,l}}\right) + \mu_{r,b,l}$
15: **end if**
16: **end for**
17: **return** S^{wiener}
18: **end procedure**

in line 3, which estimates the local mean, is of the order $\mathcal{O}(N_p \times L^2)$ as is the case for the second **for** loop in line 6 used to calculate the local variance. The noise power estimation in line 9 and the **for** loop in line 10 are of complexity $\mathcal{O}(N_p \times L)$ each. This algorithm exhibits an overall complexity of the order $\mathcal{O}(N_p \times L^2)$.

3.2.2 Wavelet denoising implementation and complexity

The 3D wavelet denoising method has been separated into two sequential wavelet denoising problems. The first carries out scan by scan 2D spatial wavelet denoising, while the second executes 1D time domain wavelet denoising for each range-bearing cell sequence. This method is presented in Algorithm 2. General wavelet methods have a complexity of order $\mathcal{O}(N \log N)$,

Algorithm 2 3D Wavelet denoising

procedure WAVELET(signal S, 1D wavelet $W^{1\mathrm{D}}$, 2D wavelet $W^{2\mathrm{D}}$, noise variance σ^2)
  for all $S(l) \in S$ **do** ▷ 2D spatial wavelet denoising
    $k_l \leftarrow \left\lfloor \log_2 \left(\frac{S(l) \text{ length}}{W^{2\mathrm{D}} \text{ length}} \right) \right\rfloor$
    $S^{\mathrm{dwt2D}} \leftarrow \mathrm{wavedec2}\big(S(l), W^{2\mathrm{D}}, k_l\big)$
    $k_{\mathrm{th}} \leftarrow \sigma^2 \cdot \sqrt{2 \log_2 \Big(\mathrm{size}\big(S(l)\big)\Big)}$
    $S_{\mathrm{th}}^{\mathrm{dwt2D}} \leftarrow \mathrm{soft\_thresh}\big(S^{\mathrm{dwt2D}}, k_{\mathrm{th}}\big)$
    $S^{\mathrm{den}}(l) \leftarrow \mathrm{waverec2}\big(S_{\mathrm{th}}^{\mathrm{dwt2D}}, W^{2\mathrm{D}}\big)$
  end for
  for all $S_{r,b}^{den} \in S^{\mathrm{den}}$ **do** ▷ 1D temporal denoising
    $S^{\mathrm{dwt1D}} \leftarrow \mathrm{wavedec}\big(S_{r,b}^{\mathrm{den}}, W^{1\mathrm{D}}\big)$
    $k_{\mathrm{th}} \leftarrow \sigma^2 \cdot \sqrt{2 \log_2 \Big(\mathrm{length}\big(S_{r,b}^{\mathrm{den}}\big)\Big)}$
    $S_{\mathrm{th}}^{\mathrm{dwt1D}} \leftarrow \mathrm{soft\_thresh}\big(S^{\mathrm{dwt1D}}, k_{\mathrm{th}}\big)$
    $S_{r,b}^{\mathrm{wavelet}} \leftarrow \mathrm{waverec}\big(S_{\mathrm{th}}^{\mathrm{dwt1D}}, W^{1\mathrm{D}}\big)$
  end for
  return S^{wavelet}
end procedure

where N corresponds to the data size (total number of pixels for 2D, or signal length for the 1D case). Therefore, the first part is of the order $\mathcal{O}(L \times N_p \log N_p)$ while the second part presents a complexity of the order $\mathcal{O}(N_p \times L \log L)$. Hence the total complexity of the wavelet method is of the order $\mathcal{O}(L \times N_p \log N_p + N_p \times L \log L)$ which reduces to $\mathcal{O}\big((N_p \times L) \log(N_p \times L)\big)$.

3.2.3 BINR implementation and complexity

The optimum parameters required for Binary Integration, $M_{\text{opt}}^{\text{BI}}$ and $P_{fa}^{\text{CA-CFAR}}$, can be calculated off-line and they are not included in the algorithm, but given as inputs. Also, the BI probability of detection, which is dependent on $M_{\text{opt}}^{\text{BI}}$ and L, can also be defined offline, and the corresponding polynomial in $P_{D_{r,b,l}}^{\text{CA-CFAR}}$ is therefore, pre-calculated. The first **for** loop in

Algorithm 3 Binary Integration Noise Reduction

1: **procedure** BINR(signal S, prob. of false alarm $P_{fa}^{\text{CA-CFAR}}$, BI parameter M_{opt}^{BI}, recursive averaging parameter α_d, noise subtraction parameters c, d)

2: $\tau \leftarrow 2W\left(\left(P_{fa}^{\text{CA-CFAR}}\right)^{-\frac{1}{2W}} - 1\right)$

3: **for all** $S_{r,b,l} \in S$ **do**

▷ Estimate SNP

4: $\widehat{\eta}_{r,b,l}^{\text{SNP}} \leftarrow \dfrac{S_{r,b,l}}{\frac{1}{2W}\sum_{j=q-W}^{q+W} S_{j,b,l} \text{ (for } j \neq q)}$

▷ Calculate CA-CFAR P_D

5: $P_{D_{r,b,l}}^{\text{CA-CFAR}} \leftarrow \left[1 + \frac{\tau}{2W}\left(\frac{1}{1+\widehat{\eta}_{r,b,l}^{\text{SNP}}}\right)\right]^{-2W}$

▷ Calculate BI P_D

6: $P_{D_{r,b,l}}^{BI} \leftarrow \sum_{l=M_{opt}^{\text{BI}}}^{L} \frac{L!}{l!(L-l)!}\left(P_{D_{r,b,l}}^{\text{CA-CFAR}}\right)^{l}\left(1 - P_{D_{r,b,l}}^{\text{CA-CFAR}}\right)^{L-1}$

7: **end for**

8: **for all** $S(l) \in S$ **do**

▷ Calculate $\tilde{\alpha}_d$

9: $\tilde{\alpha}_d(l) \leftarrow \alpha_d + \big(1 - \alpha_d\big)P_D^{\text{BI}}(l)$

▷ Update noise power estimate

10: $\Sigma_n(l) \leftarrow \tilde{\alpha}_d\Sigma_n(l-1) + (1 - \tilde{\alpha}_d)S(l)$

▷ Apply noise subtraction

11: **if** $S(l) > c\Sigma_n(l)$ **then**

12: $S_{lin}^{\text{BINR}}(l) \leftarrow S(l) - c\Sigma_n(l)$

13: **else**

14: $S_{lin}^{\text{BINR}}(l) \leftarrow d \times \Sigma_n(l)$

15: **end if**

16: **end for**

17: **return** S_{lin}^{BINR}

18: **end procedure**

line 3 has a complexity of the order of $\mathcal{O}(N_p \times L)$. The second **for** loop in line 8 involves op-

erations on each pixel for each observation, which yields a complexity of the order $\mathcal{O}(N_p \times L)$. The combination of the CA-CFAR, BI and noise subtraction parts of the algorithm yields a complexity of the order $\mathcal{O}(N_p \times L)$.

In summary, with respect to observations L, the Wiener filter algorithm has the highest complexity (quadratic), followed by the linearithmic[1] complexity of the Wavelet approach, while the binary integration noise reduction's linear complexity makes it the least complex. On the other hand, the Wiener filter and the binary integration noise reduction are the least complex (linear time) with respect to the data size N_v.

3.2.4 Experimental computational time results

The computational time used by the different algorithms, plotted against observation number L, is shown in Figure 3.3. The results are consistent with the analysis presented in the previous sections. The Wiener filter's complexity grows approximately quadratically with L, while the wavelet exhibits linearithmic complexity. A linear time complexity is achieved by the BINR method.

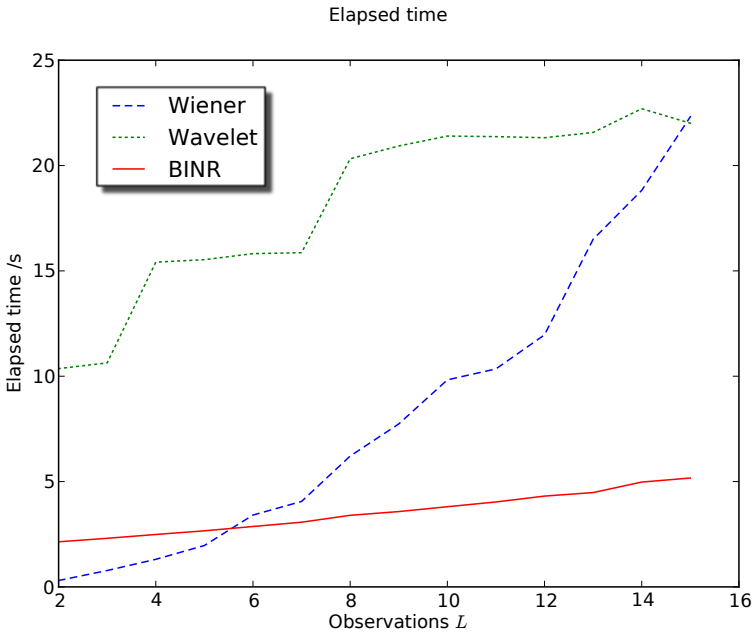

FIGURE 3.3: Elapsed computational time measurement.

[1] A *linearithmic* function is of the form $n \log n$. An algorithm with a time complexity of the order $\mathcal{O}(n \log n)$ is said to run in *linearithmic* time.

L	2	5	10	15	20
M^{BI}	1	2	4	5	7

TABLE 3.1: Optimal M^{BI} parameter for different number of observations L.

3.3 Noise reduction and detection in scanning radar data

An experimental radar data set, captured[2] in a public park in Santiago, is used to test the noise reduction schemes.

The data set contains scans from 12 different locations along a line in the main paved track of Parque O'Higgins in Santiago. At each location 160 or more scans were obtained with a scanning radar. The ground truth was labelled by hand using satellite images from Google Earth and Bing services, and a Geographic Information Software to align GPS information recorded at the targets using a hand-held GPS and by the sensor which was mounted on a Husky robot with onboard GPS. The results shown in the following sections correspond to the location where the ground truth labels exhibited a lower position error.

3.3.1 Noise reduction in scanning radar data

Noise values in real radar data do not conform to perfect Gaussian or exponential distributions, as assumed by the noise reduction methods, which impairs their performance.

An analysis of the noise reduction methods considering the park environment now follows. Although the methods were applied to the B-Scope radar data (range vs. bearing), the results are shown in *plan position indicator* (PPI) form for clearer visualization. The test environment is shown in Figure 3.4.

The area corresponds to a main paved track approximately 65 m wide. On the sides of the track there are lamp posts and some trees. There are also fences and concrete walls. The radar was located in the track.

The CA-CFAR window size was 9 bins in the bearing direction and 7 bins in the range direction. The guard cells window size was 5 in the bearing dimension and 3 bins in the range direction. These parameters were found suitable, on preliminary experiments, for detecting the lamp posts and trees surrounding the radar, by considering the power spread this features present in the acquired data.

The BI false alarm rate used was 1×10^{-6}. The optimal M^{BI} has been previously obtained for different L values. Some of the values are listed in Table 3.1 The results presented corre-

[2] using an Acumine 94 GHz, scanning radar [WCBS+06a].

FIGURE 3.4: Park environment where radar data was captured (obtained from Google Earth).

spond to $L = 20$ observations. For the noise subtraction algorithm, the chosen parameters were $\alpha_d = 0.9$, $c = 50.0$ and $d = 0.1$. A high α_d ensures that the previous value of the noise estimate has more weight than the new observation which is desirable given the high amount of noise present in radar data. Parameter c, controlling over-subtraction, was selected by testing different values between 10 and 100. Similarly, spectral floor parameter d was tested for different values between 0.05 to 0.5, with the chosen value yielding good results in the reduction of the broadband noise.

The Wiener support region was 3 bins in the range and bearing direction, which is the expected power spread of the targets of interest.

In the case of the spatial (2D) wavelet denoising, the Daubechies 3 wavelet function was used, while the Haar wavelet was selected for the 1D (temporal) dimension.

The noisy raw radar input data from the park is presented in Figure 3.5.

The ground truth location of lamp posts and trees are marked with green circles and a cross in their center.

Wiener filtering (Figure 3.6) exhibits a smoother noise background but the main objects identified in the scene are blurred by the filter, thus losing localization detail.

Wavelet denoising (Figure 3.7), is able to preserve the location and edges of targets. It does, however, produce several negative values in noise only sections, which are truncated to a small value to allow visualization. Nevertheless, the average noise level is reduced.

Finally, the BINR method in Figure 3.8 shows its ability to retain details as well as to

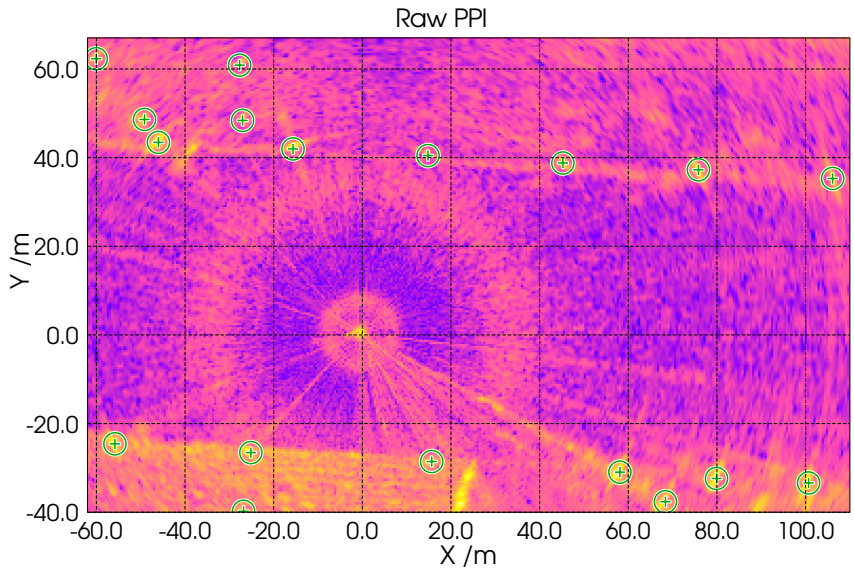

FIGURE 3.5: PPI showing noisy input data from the park environment.

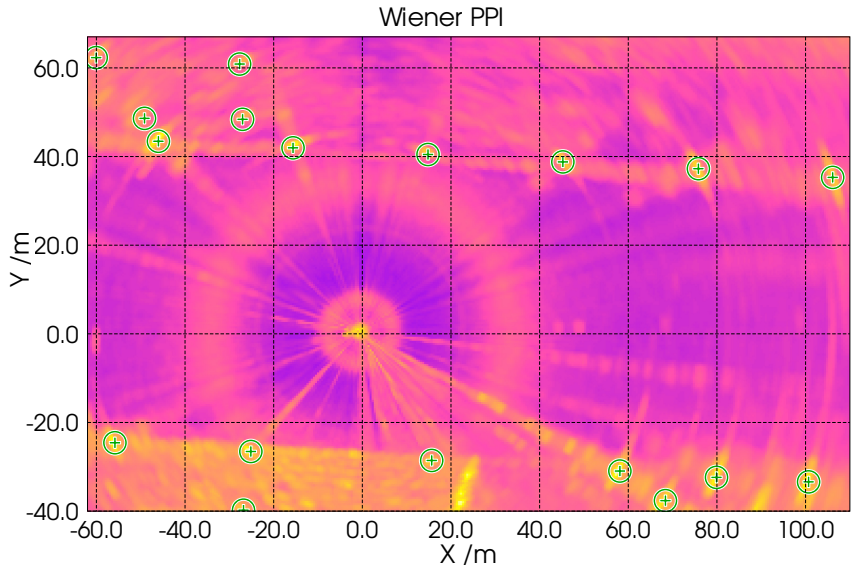

FIGURE 3.6: PPI showing Wiener filtered data from the park environment.

reduce the noise level. It can, however, be observed that some of the maximum power peaks have been reduced in magnitude (e.g. tree at $(-49.1\,\mathrm{m},\ 48.6\,\mathrm{m})$; with a raw power value of 86.02 dB and a BINR power value of 82.73 dB). This is due to the fact that at some observation l those particular targets are not detected, therefore their value is considered noise and thus subtracted from them. Note that wavelet denoising (Figure 3.7) reduces some

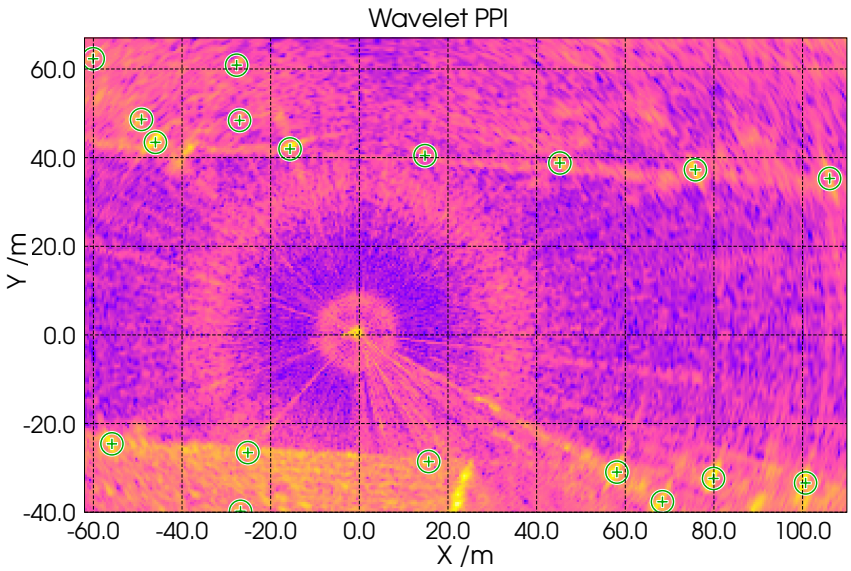

FIGURE 3.7: PPI showing Wavelet denoised data from the park environment.

noise-only areas to very low values. However, the noise background is not homogeneous, therefore, the sharp edges between noise areas near the average noise level and those greatly reduced by the wavelet method can yield several false detections as will be shown, after applying the CA-CFAR detector.

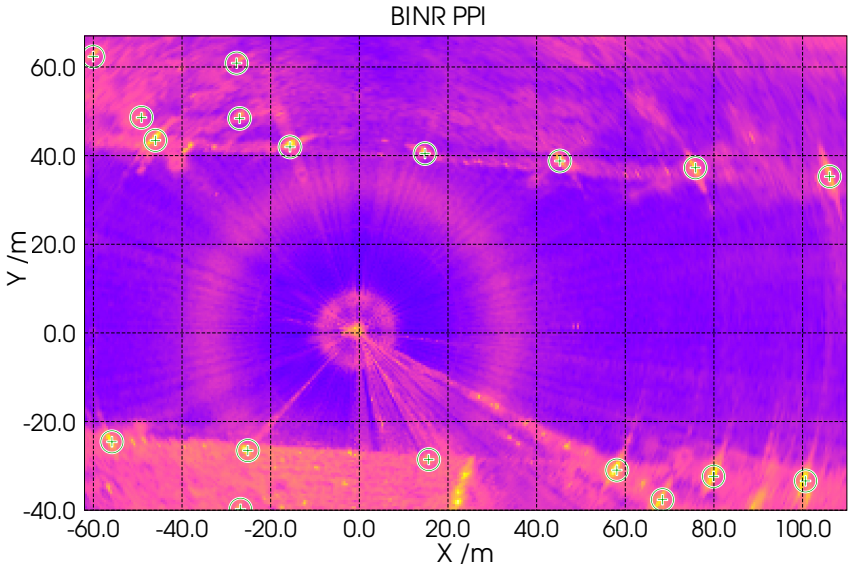

FIGURE 3.8: PPI showing BINR data from car park environment.

Figure 3.9 shows mean noise power values from each method, in an area which is known to contain no targets.

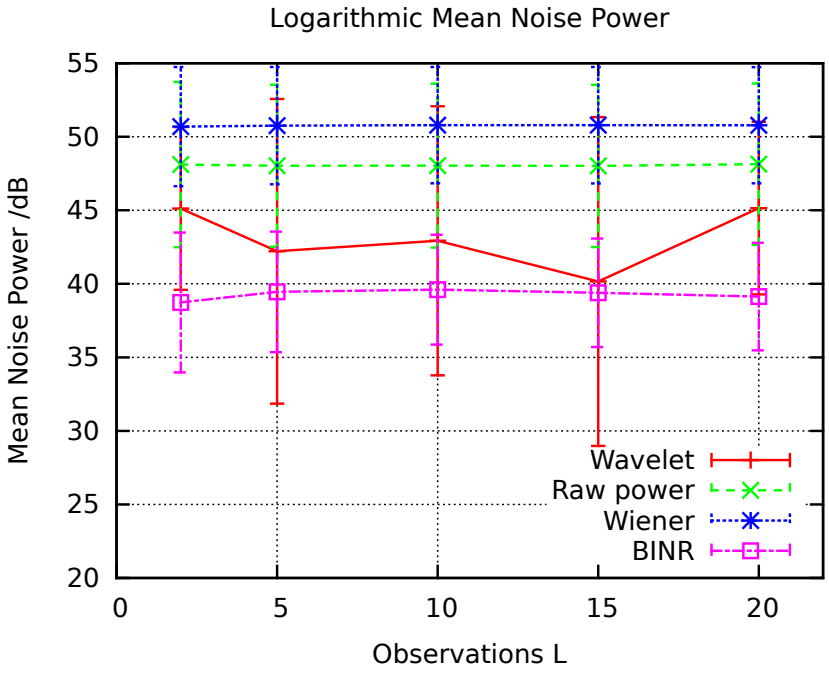

FIGURE 3.9: Noise mean values and variances from a noise only area.

The BINR method exhibits the the lowest mean noise power for all L values. Its variance, on the other hand, is higher than that of the Wiener for small L values, but as more observations are included, BINR achieves also the lowest noise variance. The Wiener filter mean noise power stays above that of the raw data, but keeps the variance at a low value. The Wavelet denoising method is able to reduce the mean noise power lower than the raw data but is not able to reach the value of the BINR method. The Wavelet's variance remains high and oscillates due to the ripple effect mentioned before. The noise assumptions which form the basis of all three noise reduction techniques are violated in practice. In particular, the Wiener filter is not optimal for non-Gaussian noise distributions. In the case of the wavelet method, the universal threshold is not able to correctly estimate a noise threshold to separate the noise and information based wavelet coefficients. Furthermore, noise information is no longer spread homogeneously across all wavelet coefficients. Likewise, the CFAR method used in the BINR yields a higher false alarm rate than expected, since exponential noise is assumed, as will be shown in the next section.

3.3.2 Target detection in scanning radar data

The CA-CFAR detector is applied to the reduced noise data, in order to demonstrate the usefulness of each reduction method. In this case a more relaxed CA-CFAR probability of false alarm is applied ($P_{fa} = 1 \times 10^{-3}$), which reduces $\tau^{\text{CA-CFAR}}$ and increments the $P_D^{\text{CA-CFAR}}$ (see Equations (2.3) and (2.2)). This is important as all reduction methods output target power which is lower than the raw data. Also, this is possible since the reduced noise data is expected to yield a lower false alarm rate.

Figure 3.10 shows the result of the detector applied to the raw noisy input data. The detector itself is able to reduce false alarms to some extent in the raw data. Again, the green circles in the figure denote the ground truth location of some relevant targets (trees and lamp posts at the side of the track) obtained using multiple scans from a laser scanner. Red square markers are used to show missed detections, while magenta coloured diamonds are used to mark false alarms.

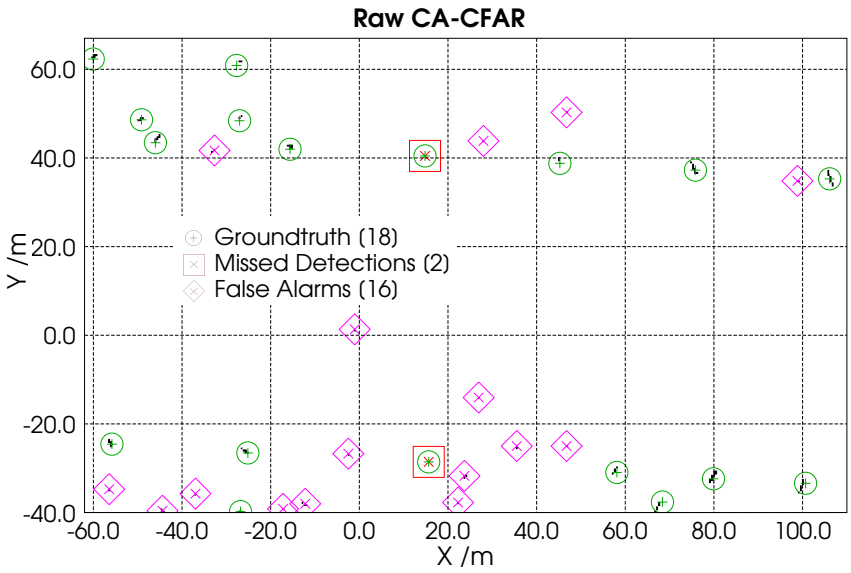

FIGURE 3.10: CA-CFAR PPI showing noisy input data from park environment.

The Wiener filter (Figure 3.11), interestingly, shows a remarkable reduction of background noise, but its blurring effects present some problems. In particular, small and medium size targets with low SNP are blurred in such a way that their size is reduced and even completely removed from the CA-CFAR detector output. On the other hand, high SNP targets become increased in size (more pixels surrounding the landmark are marked as detections), for instance, the tree at coordinates ($-27.0\,\text{m}$, $48.4\,\text{m}$). In general, edge details are lost.

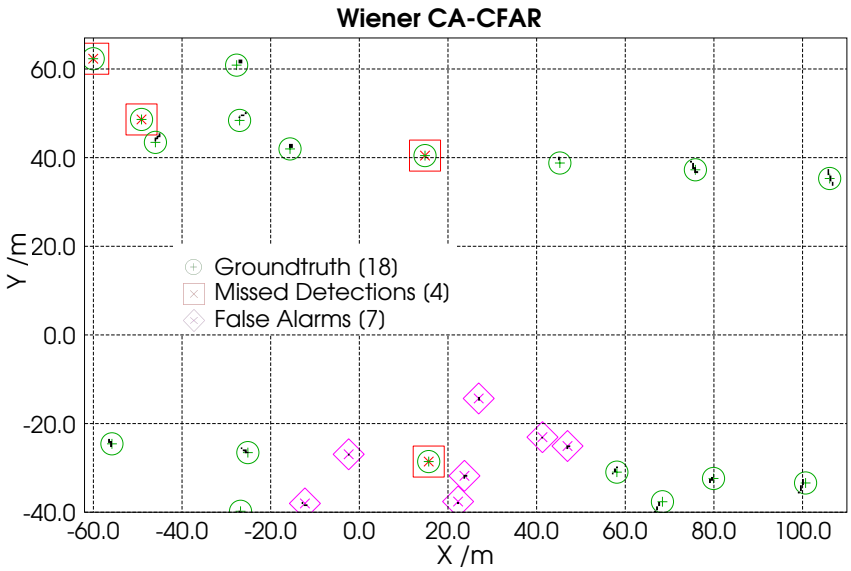

FIGURE 3.11: CA-CFAR PPI showing Wiener filtered data from park environment.

Figure 3.12 presents the results of applying the detector to the wavelet denoised data. The method, is able to keep feature details but it shows more false alarms than expected, especially in broad "noise-only" areas. This is due to the noise background not being homogeneous, which leads to the ripple effect of the chosen wavelet function, and the universal threshold not being able to correctly discriminate between information and noise wavelet coefficients.

The BINR method (Figure 3.13), is able to preserve localization details, e.g. targets at $(-25.1\,\mathrm{m}, -26.6\,\mathrm{m})$, $(68.4\,\mathrm{m}, -37.6\,\mathrm{m})$ and $(100.7\,\mathrm{m}, -33.4\,\mathrm{m})$, as well as reduce the number of false alarms when compared to the detector applied to raw data and the other noise reduction methods.

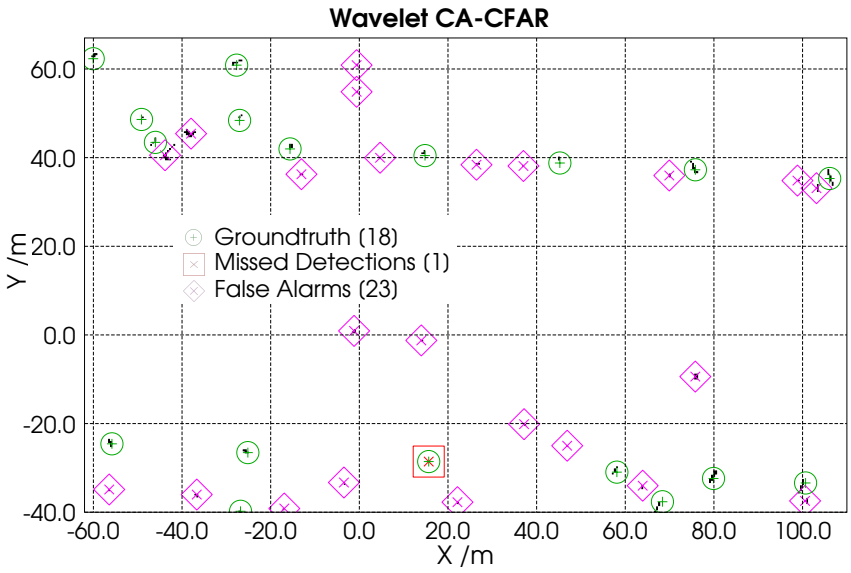

FIGURE 3.12: CA-CFAR PPI showing Wavelet denoised data from park environment.

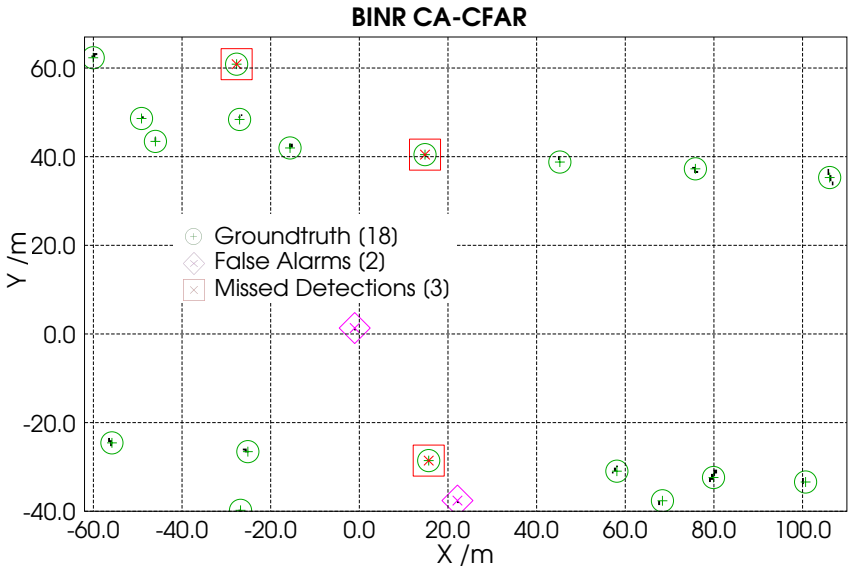

FIGURE 3.13: CA-CFAR PPI showing BINR data from car park environment.

A posteriori detection and false alarm rates can be derived from the results. Detection rate, R_D, is obtained by dividing the number of correctly detected ground truth targets, T_D, and the total number of ground truth targets, T_G, i.e. $R_D = T_d/T_G$. The False alarm rate, R_{fa}, on the other hand, is calculated by dividing the number of pixels corresponding to false alarms, T_{fa}, and the sum of pixels correctly identified as noise, T_N, plus the false alarm

Method	T_D	T_G	T_{fa}	T_N	R_D	R_{fa} /10^{-3}
Raw	16	18	23	81919	0.89	0.2807
Wiener	14	18	14	81943	0.78	0.1708
Wavelet	17	18	88	81712	0.94	1.0758
BINR	15	18	2	81977	0.83	0.0244

TABLE 3.2: A posteriori detection (R_D) and false alarm (R_{fa}) rates.

pixels, i.e. $R_{fa} = T_{fa}/(T_{fa} + T_N)$.

Table 3.2 summarizes the detection and false alarm rates calculation. Note that the area, where the experiment was carried out, exhibits a lower false alarm rate than expected (0.281×10^{-3}), with a detection rate of 89%. The Wiener filter yields a lower false alarm rate (0.171×10^{-3}) but in the same order of magnitude. The detection rate is also lower (78%) than in the raw image. Wavelet denoising's detection rate is higher (94%), but its false alarm rate is higher in an order of magnitude (1.076×10^{-3}). BINR is able to reduce the false alarm rate in an order of magnitude (0.024×10^{-3}, with a detection rate (83%) higher than the Wiener, but lower than the raw data. From this result, it can be seen that using BINR it is safe to increase the $P_{fa}^{\text{CA-CFAR}}$ value in order to get a higher detection rate while still achieving a low false alarm rate.

The metric presented in Section 2.4 was applied to the CA-CFAR output of the reduced noise images to quantify the performance of each method. The results for different L values are shown in Table 3.3. Because of the range and angular resolution of the radar used in this experiment, for each ground truth landmark, more than one point is marked as a detection. This increases the error measured by the metric because extra points are considered false alarms. Thus, the metric output starts to converge to the cut-off parameter, as the second term to the right of Equation (2.18) dominates. A method to reduce this effect is to combine multiple detections in close proximity using a clustering method, e.g. *connected component labeling* [SS02]. A different approach, consisting of duplicating the elements of the ground truth set, was used in this implementation, which effectively reduces the effect described above, without altering the results. For all L values, the BINR methods yields a smaller value of the metric meaning that the produced map is closer to the ground truth than that generated by the raw data an the other methods. The Wiener filter is also capable of producing a map with a lower error than the raw data. On the other hand, the Wavelet denoising shows a poor performance due to the false alarms in the noise background caused by the ripple effect which affects the Wavelet transform.

L	Raw power	Wiener	Wavelet	BINR
2	1.6747	1.6225	1.6565	1.5401
5	1.7055	1.6174	1.8435	1.5230
10	1.6770	1.6377	1.7328	1.4486
15	1.7013	1.6230	1.8673	1.4365
20	1.6796	1.6025	1.7878	1.5114

TABLE 3.3: OSPA metric applied to each method.

3.4 Noise reduction and detection in SAR data

Detection and noise reduction methods in radar are not only used in classical A-Scopes, B-Scopes and PPIs, they can also be used in other forms of radar data such as SAR images [dBG01, GLZ+09].

SAR images, being constructed in a fundamentally different way than the classical radar images, are affected by noise in a different way. In SAR images, noise and clutter are usually modeled by a Weibull or K distribution. Also, the effect of multiplicative speckle noise in SAR images is higher than in other forms of radar data. Under these conditions, the Ordered Statistics (OS) CFAR detection method has proven to be effective when applied to SAR images [KC94].

In this section the results of using the BINR method on a set of SAR images obtained from NASA Jet Propulsion Lab (JPL)'s Uninhabited Aerial Vehicle SAR (UAVSAR) mission[3] are presented.

The images correspond to a location near Sacramento, CA, which covers an area of crop fields with isolated buildings in the north-most part (top) of the image and a suburban area with high density housing in the south-most part (bottom). The UAVSAR mission captured data from this location in six different times. Figure 3.14 (left) shows a Yahoo Satellite image of the area, and its corresponding SAR image (right). The area is 1.6 km in the horizontal (east-west) direction and 2.88 km in the vertical (north-south) direction.

These SAR images represent backscatterred radar power, polarized in the HH, HV and VV components. The magnitude of each component is encoded in the image's red, green and blue channels, respectively. Each component was processed independently, and the final results were encoded back in the red, green and blue channels.

BINR based on the OS-CFAR detector has been used to first reduce the noise in a series of multiple ($L = 6, M^{\mathrm{BI}} = 3$) observations of the same area. Then the OS-CFAR detector is applied to the reduced noise data to detect buildings. The OS-CFAR window size was

[3]UAVSAR data courtesy NASA/JPL-Caltech. `http://uavsar.jpl.nasa.gov/`

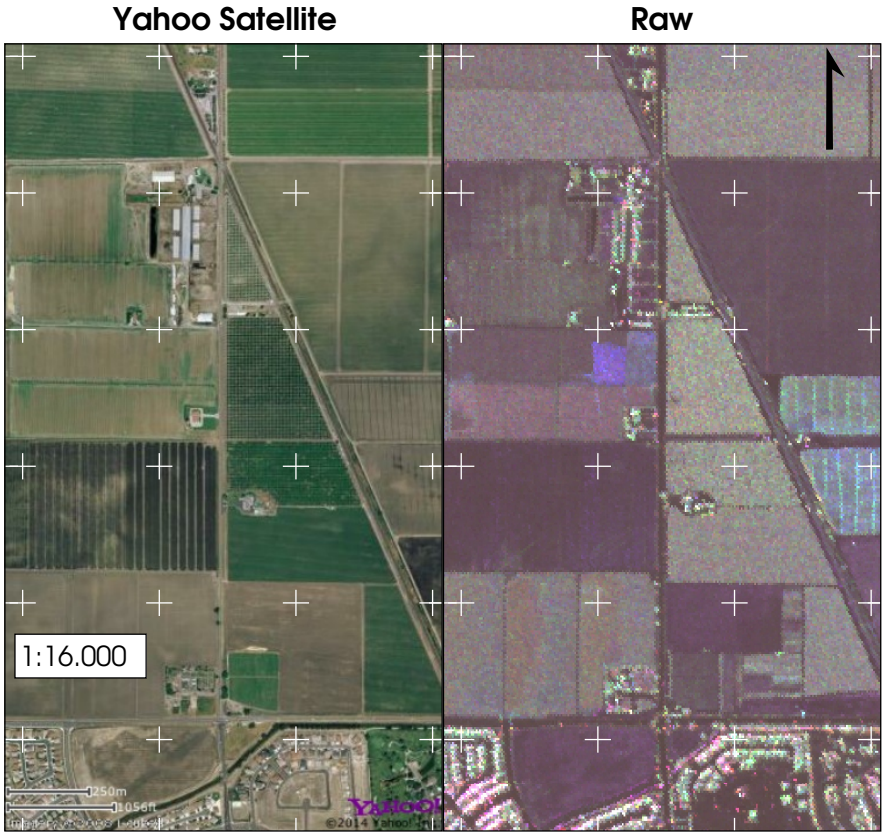

FIGURE 3.14: Yahoo Satellite view (left) and Raw SAR image of the area (right).

7 bins in the x and y coordinates, while the guard cells window size was 3 bins in both directions. The threshold constant parameter chosen was $\tau^{\text{OS-CFAR}} = 4.16707$. Finally, the noise subtraction parameters used were $\alpha_d = 0.9$, $c = 50.0$ and $d = 0.1$.

The parameters for Wavelet denoising and Wiener filtering were the same as those used in the experimental data set presented in Section 3.3.

All three components have been processed. Buildings, in general, reflect radar waves similarly in all polarizations while vegetation and other terrain considered clutter, in this case, usually exhibit different back-scatter intensity at the different polarizations.

The raw power (left) and the output of the OS-CFAR detector are shown in Figure 3.15. In the OS-CFAR image, the red, green and blue pixels corresponds to detections in the HH, HV and VV polarizations, respectively. Cyan, magenta and yellow pixels represent detections in the respective combinations of two polarizations, while white pixels represent detections in all three polarizations. Buildings appear in the OS-CFAR image with with pixels (detections in all polarizations), while parks and crop fields present detections in single polarizations or no detection at all.

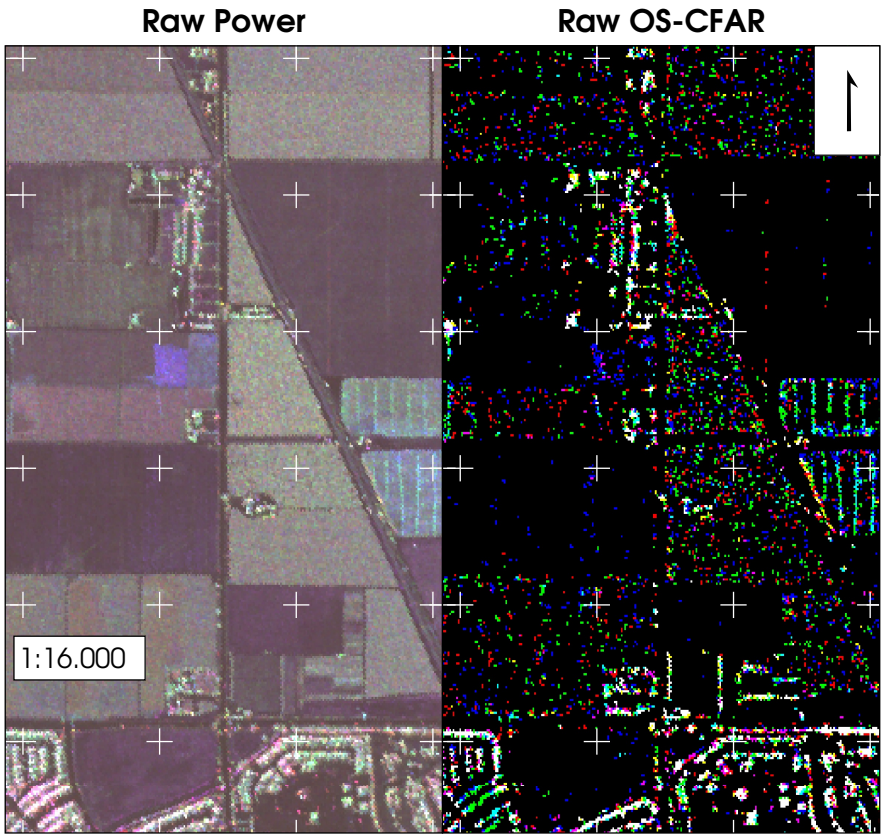

FIGURE 3.15: Raw power SAR image (left) and OS-CFAR applied to it (right).

Wiener filtering results are presented in Figure 3.16. The reduced noise image looks blurred, as expected from the Wiener filter, and the OS-CFAR detector is unable to detect buildings from the crop fields.

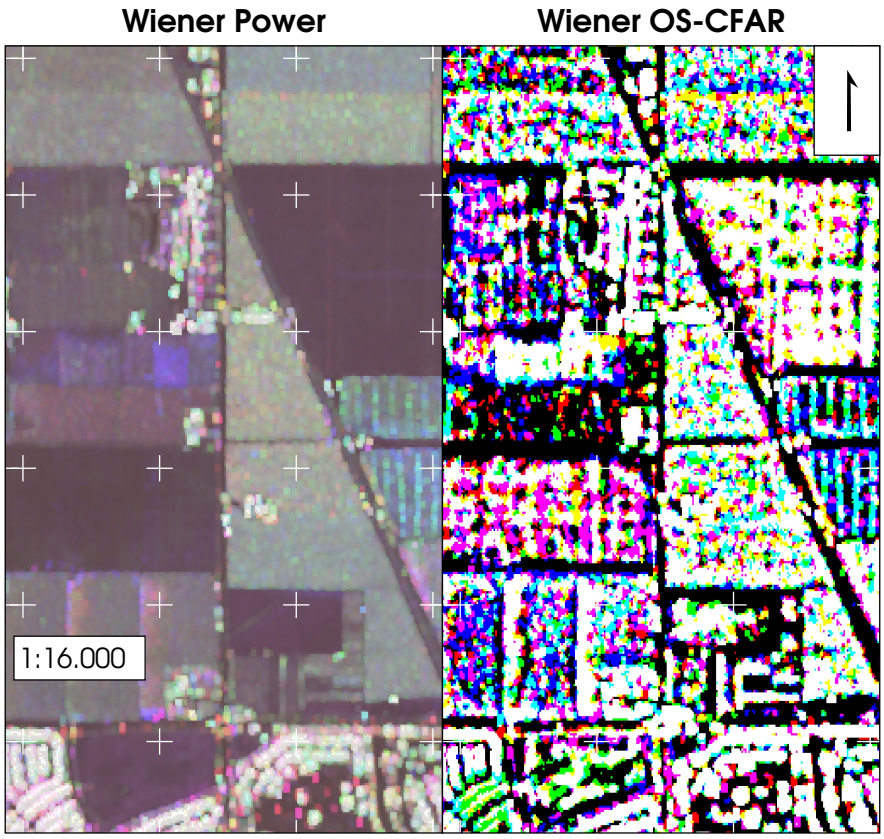

FIGURE 3.16: Wiener reduced noise SAR image (left) and OS-CFAR applied to it (right).

The wavelet denoising output is shown in Figure 3.17. The power image shows darker colours meaning that the average noise power has been reduced, but several areas present a high variance, particularly in the crop fields. The OS-CFAR output confirm this, and the detector is unable to detect building structures.

Figure 3.18 corresponds to BINR output. It can be observed that the areas corresponding to crop fields appear smoothed when compared to the raw image. It can also be observed that the number of detections in single polarizations, mostly located in areas corresponding to crop fields is reduced in the BINR image. On the other hand, most pixels corresponding to building like structures are preserved.

In this data set it is not possible to apply the OSPA metric as the real ground truth is unavailable. An analysis on the noise statistics in an area with no targets, as was carried out in the park data set, quantifies the performance of the noise reduction methods. Table 3.4 shows the mean noise power and variance per polarization channel. It can be observed that the Wavelet presents the lowest mean noise power, but at the same time it yields a very high variance. The Wiener and BINR methods keep a variance similar to that of the original data, but BINR is able to reduce the mean noise level considerably, while the Wiener filter noise

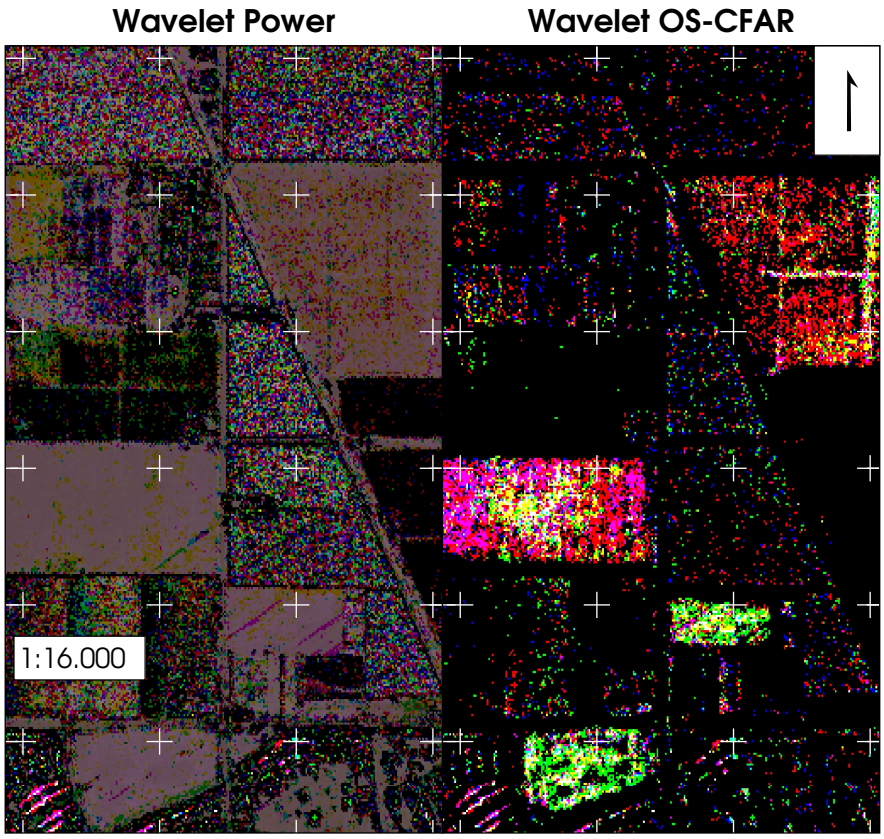

FIGURE 3.17: Wavelet reduced noise SAR image (left) and OS-CFAR applied to it (right).

	Raw		Wiener		Wavelet		BINR	
Pol.	Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.
HH	24.96	160.85	27.58	159.20	-13.64	1434.88	14.38	166.84
HV	29.14	350.91	36.59	437.78	-9.62	916.63	20.30	383.10
VV	28.62	169.56	32.74	180.68	-9.75	1042.87	19.26	204.06

TABLE 3.4: Mean noise power and variance in noise only area in dB.

level stays at a similar or higher value than the raw data.

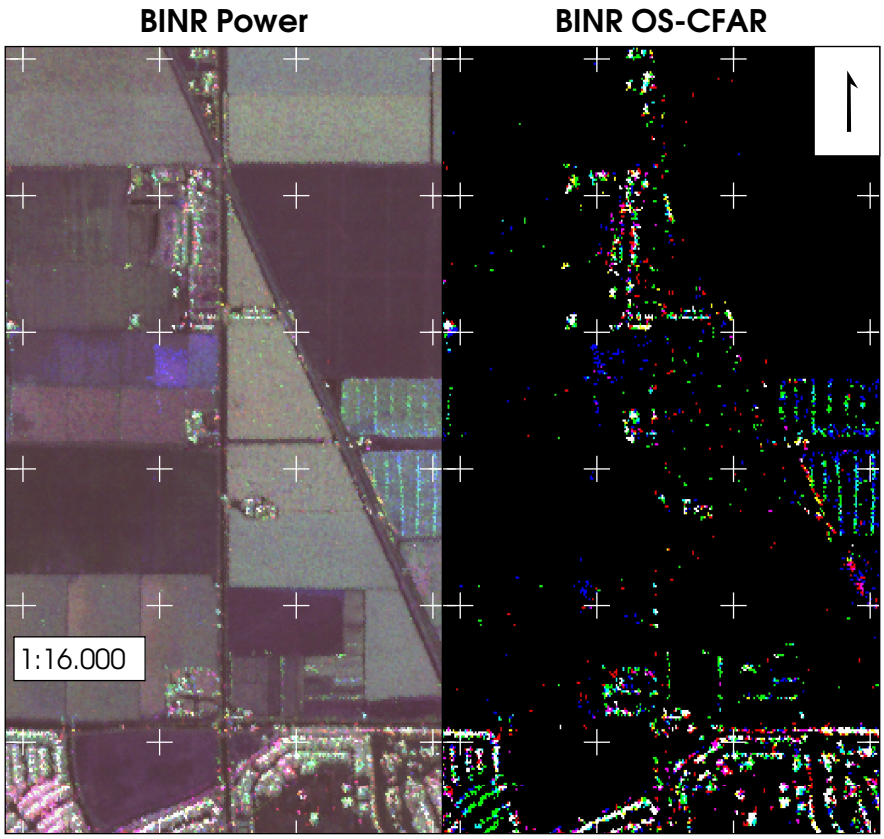

FIGURE 3.18: BINR SAR image and OS-CFAR applied to it.

Chapter 4

Detecting ladar features with a Constant False Alarm Rate in 3D point cloud data

4.1 Introduction

In several applications, when there are multiple observations (or measurements) of the same object(s), each captured into a single data set, control points are needed to identify the same part of the object(s) in the different data sets. These applications include image alignment, tracking moving objects in a scene, ladar surveying, among others and they can belong to research fields as diverse as computer vision, medical imaging, augmented reality, vibrational analysis or robot mapping. Regardless of the application, the general procedure to obtain such control points is similar. This involves the detection of some identifiable features of the objects in each data set and matching the detected features of the corresponding objects across the different data sets. Depending on the application, the features and methods of extracting them can vary widely. If the measured objects exhibit some unique characteristic which can be used for identification, this feature is usually calculated or "extracted" directly from the data set. These features are referred to as *natural* features and some examples of this approach include the well known SIFT [Low04], SURF [BETVG08], ORB [RRKB11] and surface normal features used in 2D or 3D [MN03] image processing. Another example of natural features occurs in the analysis of the waves in an electrocardiogram (ECG). In particular, the so-called R peak (specific part of the QRS complex in the cardiac cycle) is used to identify the cycle and to measure the cardiac frequency (see Fig. 4.1).

When the objects lack an identifiable characteristic, or the ability to extract this type of feature is not accurate enough, artificial markers are added to the observed objects. Then, the markers provide the needed features in the data set, which can be measured depending on the

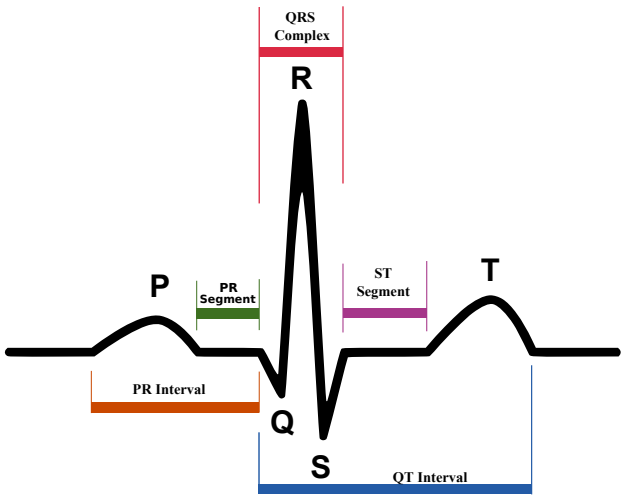

FIGURE 4.1: Schematic cardiac cycle showing the R peak and other wave complexes.

nature of the marker. For instance, the widely used QR (Quick Response) codes incorporate black squares as position and alignment markers (see Fig. 4.2). Depending on the marker's distinguishable information, which can be shape, reflectivity, colour, pattern, texture or local gradient, different methods for detecting these features should be used. Often, combinations of these characteristics are exploited simultaneously.

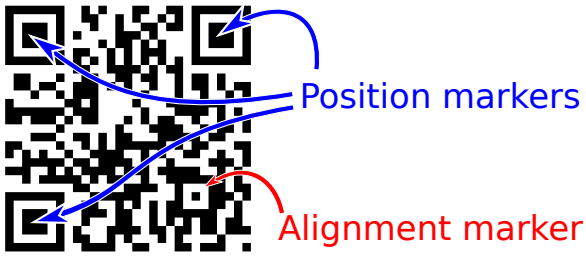

FIGURE 4.2: Sample QR code with position and alignment markers.

The detection of both, natural or artificial features can be used for the fundamental task of transforming different sets of data into a common coordinate system, called *registration* [Bro92, Gos05]. When a high level of accuracy is required, the use of *artificial* landmarks is unavoidable [SWB02].

In the particular case of ladar surveying, retro-reflective targets are used as artificial markers, due to the great variability in the measured data, which would make it difficult to extract natural features in some situations. These targets provide a high intensity measurement, which usually stands out from the rest of the data. They also provide a distin-

guishable shape. Commercially available 3D ladar scanning solutions include sets of artificial landmarks, such as those shown in Figure 4.3 and Figure 4.4.

FIGURE 4.3: Riegl Retroreflector targets

FIGURE 4.4: Riegl Retroreflector targets (schematic)

The detection of these markers can be manual, interactive or automatic. Manual methods require a person to identify and tag the markers, whereas interactive methods require human intervention to guide the process. These two methods are both time-consuming and a potential source of error [KHW+03]. Fully automated methods, on the other hand, do not require any intervention from the user. These are usually desirable as reliability can be increased or quantified, while time and costs can be reduced. Most automated feature extractors for 3D data, such as that obtained from ladar scanners, are currently based on 2D algorithms developed for camera images, and thus, usually discard the range information available from a 3D ladar. Further, current feature detection methods do not take into consideration feature

probabilities of detection or false alarm, thus generating a non-negligible number of weak features, usually leading to false matches.

As described in section 2.3, there are a number of scenarios in which artificial markers are preferred to natural features. Highly dynamic environments, environments with a small number or low quality natural features, when high accuracy and reliability is required or when the high computational complexity of extracting natural features is not affordable are situations in which artificial landmarks are either unavoidable or perform better than natural features [SKC13, BPR12, Fia10, BGG$^+$04, NHY$^+$00, TS13, NS12, ROSF11].

Some standard automated procedures to detect artificial markers are [Kal10]:

Template matching: This is a technique used to find small specific parts of an image which match a template image or computer generated model. It is commonly found when the shape of the template is known, such as a circle with a known radius or a coded target.

The basic method of template matching uses a convolution mask (containing the template) which is scanned over the searched image to detect similar features in it. The convolution output will be highest at places where the image structure matches the mask structure.

The drawback of template matching is that every position in the search image has to be considered to find matching features, this makes the technique computationally inefficient.

Edge detection and ellipse fitting: In this method, edges are detected and used as image structures for object and pattern recognition, as well as image measurements. All objects of interest stand out from the background on the basis of a characteristic change in the relevant image structure, i.e. significant changes in pixel values occurring along physical object edges. Edges can also be used to measure some characteristics of the target, such as its centre or surface area, by fitting, for example, arbitrary ellipses to the edges using a least squares method.

There are several edge detection methods, such as the Canny edge detector, which are able to detect weak or not clearly visible edges even in the presence of noise. The Canny detector consists of the following steps:

1. Image smoothing is used to reduce noise from the original image.
2. An image gradient is calculated to highlight regions with high spatial partial derivatives.
3. An edge direction is computed from the gradient values in the x and y directions.

4. An edge direction is related to the eight directions, which can be traced in the image (connected pixels).
5. Non maximum suppression is applied by tracing along the edge directions and any pixel value that is not considered to be an edge is removed.
6. Breaks in the edges are removed by using a threshold. If the difference is below the threshold the broken edges are joined, otherwise they remain disconnected.

Image segmentation This method partitions an image into non-overlapping regions and boundaries, so that pixels belonging to a region have uniform values. Several segmentation methods are available. The most common and simplest method used is based on thresholding. Pixels with values above the threshold are identified as target pixels, while those below it are discarded as non-target pixels. Threshold segmentation is a simple but fast method. It's main drawback is the appropriate selection of the threshold value to separate the targets from the rest of the data. Often, a second method is applied after thresholding to discard false alarms.

It will be shown that the nature of the 3D data produced by ladars can affect the performance of these techniques.

In this work, automatic methods for detecting reflective markers using received power and combined power-range information from ladar sensors is presented. These approaches are based on the adaptive threshold obtained by using CFAR (Constant False Alarm Rate) processors, which are used to segment the markers from the rest of the data. Comparisons are made with state-of-the-art detection and segmentation methods such as constant threshold, Canny Edge detection and RANSAC shape detection.

Initially, standard Cell Averaging (CA) and Ordered Statistics (OS) CFAR processors are applied to 3D laser surveying data projected into 2D from the point of view of the sensor, analogously as a map projection of the terrestrial globe. The contributions of this chapter then include:

1. The formulation of an adaptive parametric version of the CFAR processors, capable of combining power and range information available in the 2D projected data.
2. A modification of the CFAR algorithms applied directly to 3D point cloud data.

Three standard methods will be used as benchmarks: 1) constant threshold segmentation (for 3D and 2D projected data), 2) Canny edge detection (for 2D projected data only) and 3) RANSAC shape detection (for both 2D and 3D representations). All the methods are followed

by a standard connected component analysis (CCA) to identify and label the detected areas (2D data) or volumes (3D data).

In summary, the use of modified CFAR processors to detect reflective markers will be shown to exhibit a high performance in the tested environments. They present themselves as a good alternative over standard segmentation and detection methods, such as thresholding, edge detection, connectivity analysis or clustering, in the application of 3D ladar surveying.

4.2 Laser surveying in 3D

In order to reconstruct a 3D environment using laser surveying, several scans are obtained, usually from different positions. Thus, registration is required to locate all of the data into the same reference coordinate system.

The registration process flow in state-of-the-art 3D laser surveying consists of:

1. Acquiring a series of 3D laser scans.
2. Finding initial (before any iteration of the algorithm) feature or point correspondences for every scan pair.
3. Computing the relative orientation and translation between scans.

The detection methods presented in this work can also be used to obtain an accurate initial point correspondence, which is essential to reduce the overall error in the registration process.

The two main families of registration methods are feature based or point based. Both types are described in the following sections.

4.2.1 Point based registration

Point-based algorithms do not require the detection of features. Instead, they match point clouds by selecting point correspondences. Iterative closest point (ICP) matching is a well known algorithm for point-based registration (Besl and McKay [BM92b]). The concept behind this method is to automatically calculate point correspondences on the basis of the smallest distance between point clouds and minimizing the resulting Euclidean error. This iterative algorithm is designed to converge to a local minimum. Good start estimates are essential to achieve low error matching, thus ensuring the convergence to a correct minimum. The method proposed in this paper can be used to provide start estimates for an ICP based registration method.

4.2.2 Feature based registration

Feature based registration methods use features as correspondences in each scan pair to compute a transformation matrix. Most state-of-the-art registration methods, such as the Iterative closest point (ICP) method [BM92a], rely on the accuracy of the initial pose estimates, based on Global Position System (GPS) measurements [BSSC95] or local positioning using artificial landmarks as references [WTGBS08].

Feature based registration can use both natural or artificial features as described in section 2.3.

An issue with natural features based methods is that they do not take into consideration detection statistics. In many situations, the number of falsely declared detections (false alarms) is not constrained or even measured. The performance of any matching algorithm used for registration will be detrimented by the appearance of false alarms, which can be significant in environments with poor atmospheric conditions, such as in the presence of dust, snow, rain, etc.. The CFAR based methods proposed in this work make use of a consistent statistical analysis based on detection theory. This approach allows the user to preset an acceptable, desired false alarm rate, which can reduce the error in registration. Furthermore, each declared detection has an associated probability of detection which can be used as a measure of how reliable each detected target is.

Moreover, some environments lack the number or quality of natural features needed to achieve a required level of accuracy. Also, the environment might by highly dynamic, thus altering the potential natural features from scan to scan. Typically these environments include industrial settings, indoor spaces crowded with multiple moving agents, underground mines, robotic surgery rooms, and such. In these situations the use of artificial landmarks is preferable.

Most of the artificial landmarks approaches described in section 2.3 use markers specifically designed for a particular task. The methods presented in this paper use a generic, off-the-shelf marker such as the one already used in laser 3D surveying. Unlike, the detection methods usually applied in standard 3D surveying which require a special hardware and a preliminary phase to locate the markers before scanning the environment, the method in this work can detect the markers directly from the scanned data. Also, they are not limited to using laser retro reflective markers, but to any kind of marker which can yield high intensity reflections within the sensed data.

4.2.3 The 3D point cloud data and the sparsity problem

As a 3D ladar scans the environment, the space is discretised by the sensor in all 3 dimensions. Consider the laser sensor at the origin of the coordinate system. Because of the scanning movement, spherical coordinates (range r, bearing angle ϕ and elevation angle θ), are preferred to assign 3D coordinates to each measured data point. Besides the spatial coordinates, the data points can also record the reflected intensity value. Each point detected by the ladar corresponds to a particular voxel (volume element) of the environment. A typical ladar scanner does not record returned intensity for all voxels along a particular laser beam at (ϕ, θ), but only those voxels where an object reflects enough laser energy as to be detected by the sensor. The set of detected points is commonly referred to as Point Cloud Data (PCD). Thus if the spherical space (within the ladar's field of view) is considered as a 3D matrix containing cells which provide reflective intensity values, then not all matrix's cells have a value, since most of them are effectively empty, because only cells corresponding to voxels with objects detected by the ladar will hold an intensity value. Such a data matrix is considered to be sparse, in contrast to a dense matrix, when all cells have a well defined value. This is an advantage seen from the point of view of data size, because ladar sensors can achieve a very high resolution, producing millions of points in a point cloud. On the other hand, because of this difference, 3D point cloud data cannot be processed in the same way as 1D signals, 2D images, video and some 3D data which is represented by dense matrices. As will be discussed in section 4.4, data processing methods need to be adapted for direct use on PCD. Alternatively, a projected 2D image obtained from the PCD can also be used, but the sparsity is also transferred to the 2D representation, which then causes problems with 2D image processing algorithms. Also, it should be noted that as range increases, the voxel volume representing a particular point increases together with the distance between "adjacent" points (i.e. adjacent voxels).

Therefore, some of the difficulties arising from the nature of the 3D PCD are:

- When the 3D data is projected to a 2D image, targets at different ranges from the sensor appear to have different sizes. Also, if the sensor is not oriented along the line of sight of the target's surface normal, its shape will be distorted. This means that, for example, template matching techniques would need to test templates of different sizes and shapes, making the method even more computationally expensive than it already is. Also, any shape or size detector which might be applied after segmentation thresholding would also be affected by this issue.

- In either 3D or 2D projected data, some non-target objects might present a similar reflected intensity as the targets themselves, which can confuse threshold detectors.

This is equivalent to the noise/clutter distribution not being well separated from the target distribution (see Section 4.3.1). In this case, usually a second detector is used to discard false-positives, such as a shape or size detector, as mentioned above. The sparsity of the data, however, can affect the results of these methods. For instance, discontinuities in the pixels' (in 2D projected data) or points' (in 3D data) intensity values will make size or shape detectors fail. Also, edge detectors are affected by this phenomenon as false edges arise at these discontinuities, which can be misinterpreted as the real targets.

The proposed detection methods, highlighted in Section 4.3, have been modified to cope with some of the aforementioned limitations. Deeper discussion and analysis on these issues will be carried out in the following sections.

4.3 CFAR-based detector for ladar

In order to address the object detection problem required to extract features or to obtain start estimates in the scan registration methods described in Section 4.2, stochastic methods providing adaptive thresholds can be used. Constant False Alarm Rate (CFAR) processors comprise a family of such methods well known in the radar community. They are capable of performing in situations of low signal-to-noise ratio (SNR), achieving high probabilities of detection while bounding the false alarm rate. These methods have also been used for other types of data. One of these applications consists in an image processing technique to detect electrical lines in laser images based on CFAR processors [SF96]. The method is used for the detection of small co-linear obstacles in images from a laser sensor to be used in the avoidance of electrical lines and other small objects in helicopter airborne operations. The main problems associated with the detection of electrical lines are sensing them with a reasonable resolution from a fast moving platform and detecting the cables in cluttered range-intensity images. CFAR methods have also been used to detect and track dim points with low SNR in infrared image (IR) sequences as reported in [ZLS05]. The method uses a CFAR approach as a means to detect low SNR, small, targets in the IR images without prior knowledge of noise and clutter distributions. It was shown that the method exhibits a high performance in situations of target velocity mismatch and target manoeuvring, which are the main factors which cause miss detections in other comparable methods. The method pre-processes the image sequence by tracking the target before detection. Unlike the new approaches presented here, the method reported in [ZLS05] is used in moving target situations and therefore is not directly applicable to 3D laser surveying. The method uses a target model

specific to infrared moving targets defined by a 2D infrared optical blur function. On the other hand, an image pre-processing stage to remove atmospheric infrared noise is applied to the data. Then, complex background noise in a sequence of images is reduced using the top-hat transform from mathematical morphology. After the pre-processing stage, a track before detect (TBD) algorithm which analyses possible movement in four directions for each pixel is used before a CFAR processor is applied to declare a detection on a moving target.

The methods of automatically detecting reflectors presented in this work use the intensity level of laser energy reflected back by markers in a ladar image. Range information is also incorporated to calculate an adaptive threshold, motivated by the results obtained in [SF96]. The original CFAR methods are to be adapted to disambiguate the reflective markers from the rest of the objects in the ladar images. These stochastic methods are designed to keep the false alarm rate constant at a predetermined (preferably low) level, and they exhibit a relatively low computational complexity with respect to other detection methods, in an analogous way to the work presented in [ZLS05].

The proposed methods can also estimate the probability of detection for each point, particularly, those identified as a detection, thus, serving as a measure of how likely that particular point corresponds to an actual marker. This can be used by automatic registration systems to give a higher weight to the markers with a higher detection rate than to those with lower values. Otherwise, it can be used in an interactive or semi-automatic registration process to tell the user which markers are less reliable and should be carefully checked.

4.3.1 Ladar target noise distributions

Before applying the CFAR methods to the ladar data, it is necessary to check whether the assumptions for the validity of the CFAR algorithms hold, the most important of which are the noise and target distributions.

Ladar data captured in two different environments, an indoor laboratory and an underground mine, is used to determine the corresponding clutter and target distributions. The received power intensity values of points where there are no markers are used to obtain the clutter distribution (see Figures 4.5 and 4.6), while points corresponding to markers are used for the target distribution (Figure 4.7 shows a sample target distribution). Note that the reflectance values recorded by sensor used for the experiments (Riegl VZ-400) correspond to a reflected intensity relative to the intensity of a white diffuse target at the same distance, thus the reflected intensity values are unit-less. Also, the sensor stores these values in dB. The linear reflectance is converted from the logarithmic values.

For all the experiments the same type of markers was used, which correspond to Riegl's

standard markers as depicted in Figures 4.3 and 4.4.

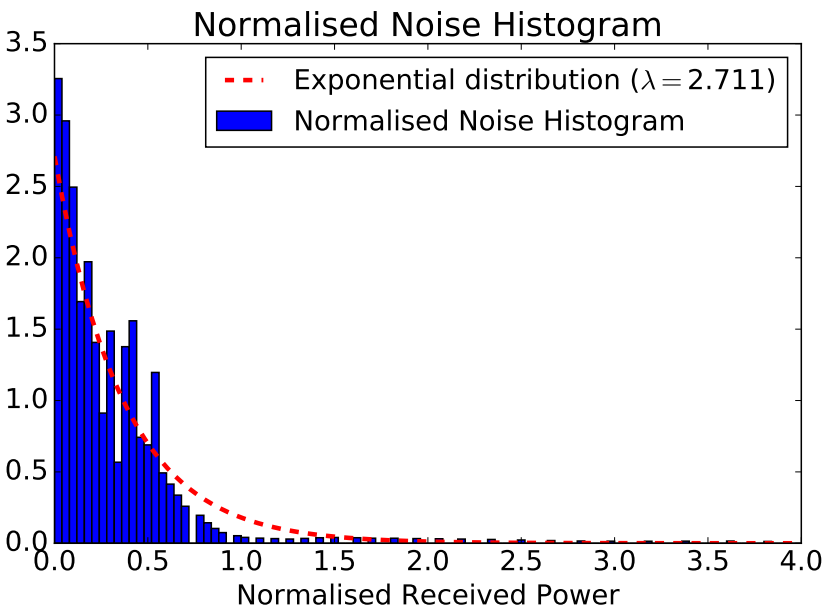

FIGURE 4.5: Ladar clutter histogram in an indoor environment.

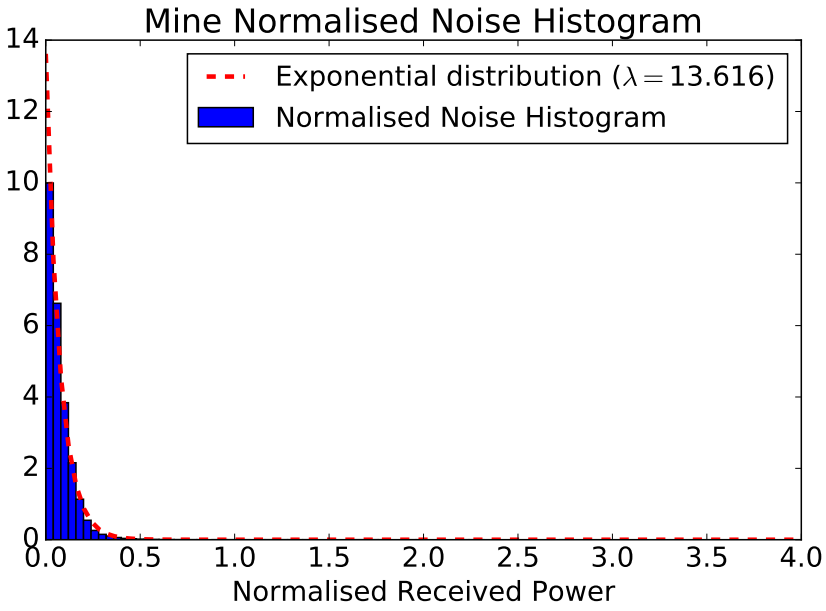

FIGURE 4.6: Ladar clutter histogram in an underground mine environment.

As seen in Figure 4.5, the clutter distribution resembles an exponential distribution. This distribution, together with the distribution for targets observed in Figure 4.7, resemble those

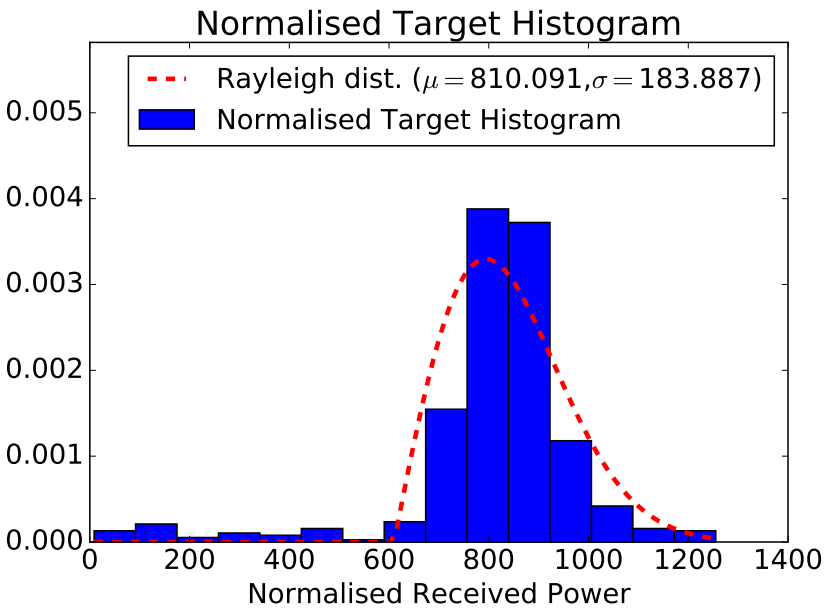

FIGURE 4.7: Target histogram for indoor environment.

reported in the radar literature [AJ12], thus the use of CFAR methods seems a sensible choice.

4.4 Adaptation of CFAR algorithms to 3D scenes

The 3D ladar sensor used in this work outputs, in spherical coordinates, a value for the reflected laser intensity received by the sensor, as well as estimates of each range r, bearing angle θ and elevation angle ϕ.

In order to use the CFAR methods with 3D ladar data, two approaches are proposed. The first one is to project the 3D ladar data onto a 2D image and then apply 2D CFAR methods. This is a common approach to visualize 3D data (e.g. cartographic maps, ladar scenes, etc). Although there are many mature, robust and fast algorithms to process 2D data, all projection techniques distort the spatial relations of objects in the original image, leading to different errors or performance detriments. Thus, a second alternative is to apply the CFAR techniques directly onto the undistorted 3D ladar data. In general, 3D algorithms are less developed than their 2D counterparts and are usually slower because of the higher amount of data.

Standard CA-CFAR, OS-CFAR, or other CFAR methods suitable for 2D data, can be directly applied to the ladar intensity image, in the same way they are used with radar data.

The standard CFAR methods will be used as benchmarks to compare the performance of the two CFAR extensions proposed in this work, both of which aim at incorporating range and reflectance information simultaneously into the CFAR analysis. The performance of the CFAR methods is expected to improve when using range information. The first one corresponds to an adaptive CFAR method which uses range information to adapt the size of the reference and guard cell windows. The second proposed variation corresponds to the direct CFAR implementation in 3D space mentioned above. These two approaches are analyzed in the following sections.

4.4.1 Projecting 3D range and intensity information into 2D spaces

In this approach the 3D ladar intensity values are projected onto a 2D image using bearing and elevation angle values. Similarly, range information can also be projected using a similar transformation. This approach has the drawback that any projection transform introduces distortions which might affect the shape of the markers and consequently detriment their detection. Also, areas of the image with no reflected data should be filled in with a "maximum range/minimum intensity" value or another specific value, to avoid false detections, which were observed in preliminary experiments. A thorough analysis of projections and effects in feature-based detectors in 3D ladar applications can be found in [HEBN13]. Two of the projections presented in the cited work will be used here. The first one is the Equirectangular projection, because it is the simplest and the most widely used projection in image processing. The other one is the Mercator projection, which is mainly used in cartography and geoinformatics, and was shown in [HEBN13] to outperform other projections for 3D laser surveying, by achieving high registration rates in different kinds of environments (indoor and outdoor) as well as for different combinations of descriptors (SIFT, SURF, ORB) and feature detectors (like SIFT, SURF, ORB, STAR and FAST), this projection also has the capability of processing a full 360° panoramic view.

In general, projecting 3D data onto a 2D image requires mapping the points (in this case spherical coordinates θ, ϕ, r) to a 2D x, y coordinate image. The *equirectangular* projection is the simplest projection and it is used in many applications to map a portion of a surface of a sphere to a flat image. In this projection, the bearing and elevation angles are directly mapped to horizontal and vertical coordinates of a grid without any transformation or scaling. As a result, vertical straight lines in the 3D space remain vertical straight lines in the 2D map, but horizontal straight lines become curves, except for the horizon. Also, poles are stretched to the entire width of the image at the top and bottom edges.

The *Mercator* projection is related to the equirectangular and to another projection called

cylindrical. It presents less distortion than its two parent projections. It also has less vertical stretching and a greater vertical field of view. Table 4.1 summarizes these two projections as described in [HEBN13]

TABLE 4.1: Summary of projections used.

Projection	f_x	f_y	hor. FoV	vert. FoV
Equirectangular	θ	ϕ	360°	180°
Mercator	θ	$\ln\left(\tan\phi + \frac{1}{\cos\phi}\right)$	360°	150°

Once intensity and range data has been projected on the respective 2D images, the standard CFAR processors are applied to obtain the adaptive threshold.

4.4.2 Adaptive Window CFAR

Intensity and range data can also be combined to improve the output of the 2D CFAR processor. This work proposes the use of range information to adjust the size of the CFAR window to match the expected size of the markers, which results in a CFAR processor with adaptive (window size, guard cells, τ) parameters. Therefore, for each cell under test (CUT), the range information is used to calculate the appropriate window and size of the guard cells in the projected bearing-elevation space. By matching the expected size of the markers into the guard cells window, the reference cells window will contain only noise or clutter values. Thus, the CFAR assumptions would be enforced and the performance of the detector should be improved. A description of this proposed method now follows.

Let $\theta_{\text{CUT}}, \phi_{\text{CUT}}, r_{\text{CUT}}$ be the CUT's coordinates in the original space, R_{gc} the radius of the target retro-reflector disk, which is used as the size of the guard cells region, and R_{rc} the radius of the desired reference cells region. Both radii are defined in Cartesian space. If θ_{gc}^{+} and θ_{gc}^{-} represents the border angles of the guard cells region in the original space, and θ_{rc}^{+} and θ_{rc}^{-} the border angles of the reference cell region in the original space, then

$$\theta_{\text{gc}}^{+} = \theta_{\text{CUT}} + R_{\text{gc}}/r_{\text{CUT}} \qquad \theta_{\text{gc}}^{-} = \theta_{\text{CUT}} - R_{\text{gc}}/r_{\text{CUT}} \tag{4.1}$$

$$\phi_{\text{gc}}^{+} = \phi_{\text{CUT}} + R_{\text{gc}}/r_{\text{CUT}} \qquad \phi_{\text{gc}}^{-} = \phi_{\text{CUT}} - R_{\text{gc}}/r_{\text{CUT}} \tag{4.2}$$

$$\theta_{\text{rc}}^{+} = \theta_{\text{CUT}} + R_{\text{rc}}/r_{\text{CUT}} \qquad \theta_{\text{rc}}^{-} = \theta_{\text{CUT}} - R_{\text{rc}}/r_{\text{CUT}} \tag{4.3}$$

$$\phi_{\text{rc}}^{+} = \phi_{\text{CUT}} + R_{\text{rc}}/r_{\text{CUT}} \qquad \phi_{\text{rc}}^{-} = \phi_{\text{CUT}} - R_{\text{rc}}/r_{\text{CUT}}. \tag{4.4}$$

Also, if $f_x(\theta, \phi)$ and $f_y(\theta, \phi)$ are the transform functions from the original to the projected space, as defined in table 4.1, then the reference and guard cell window sizes in the projected

space, represented by $(x_{\text{rc}}^{\pm}, y_{\text{rc}}^{\pm})$ and $(x_{\text{gc}}^{\pm}, y_{\text{gc}}^{\pm})$, respectively, are given by

$$x_{\text{gc}}^{+} = \max_{\theta_{\text{gc}},\phi_{\text{gc}} \in W_{\text{gc}}} (f_x(\theta_{\text{gc}}, \phi_{\text{gc}})) \tag{4.5}$$

$$x_{\text{gc}}^{-} = \min_{\theta_{\text{gc}},\phi_{\text{gc}} \in W_{\text{gc}}} (f_x(\theta_{\text{gc}}, \phi_{\text{gc}})) \tag{4.6}$$

$$y_{\text{gc}}^{+} = \max_{\theta_{\text{gc}},\phi_{\text{gc}} \in W_{\text{gc}}} (f_y(\theta_{\text{gc}}, \phi_{\text{gc}})) \tag{4.7}$$

$$y_{\text{gc}}^{-} = \min_{\theta_{\text{gc}},\phi_{\text{gc}} \in W_{\text{gc}}} (f_y(\theta_{\text{gc}}, \phi_{\text{gc}})) \tag{4.8}$$

$$x_{\text{rc}}^{+} = \max_{\theta_{\text{rc}},\phi_{\text{rc}} \in W_{\text{rc}}} (f_x(\theta_{\text{rc}}, \phi_{\text{rc}})) \tag{4.9}$$

$$x_{\text{rc}}^{-} = \min_{\theta_{\text{rc}},\phi_{\text{rc}} \in W_{\text{rc}}} (f_x(\theta_{\text{rc}}, \phi_{\text{rc}})) \tag{4.10}$$

$$y_{\text{rc}}^{+} = \max_{\theta_{\text{rc}},\phi_{\text{rc}} \in W_{\text{rc}}} (f_y(\theta_{\text{rc}}, \phi_{\text{rc}})) \tag{4.11}$$

$$y_{\text{rc}}^{-} = \min_{\theta_{\text{rc}},\phi_{\text{rc}} \in W_{\text{rc}}} (f_y(\theta_{\text{rc}}, \phi_{\text{rc}})) \tag{4.12}$$

where

$$W_{\text{gc}} = \{\theta_{\text{gc}}, \phi_{\text{gc}} | \theta_{\text{gc}}^2 + \phi_{\text{gc}}^2 \leq R_{\text{gc}}\} \text{ and} \tag{4.13}$$

$$W_{\text{rc}} = \{\theta_{\text{rc}}, \phi_{\text{rc}} | R_{\text{gc}} < \theta_{\text{rc}}^2 + \phi_{\text{rc}}^2 \leq R_{\text{rc}}\} \tag{4.14}$$

are the sets of angles corresponding to the guard and reference cell regions, respectively, defined in formal set builder notation [Ros07]. See Figure 4.8 for a graphical representation of all these parameters.

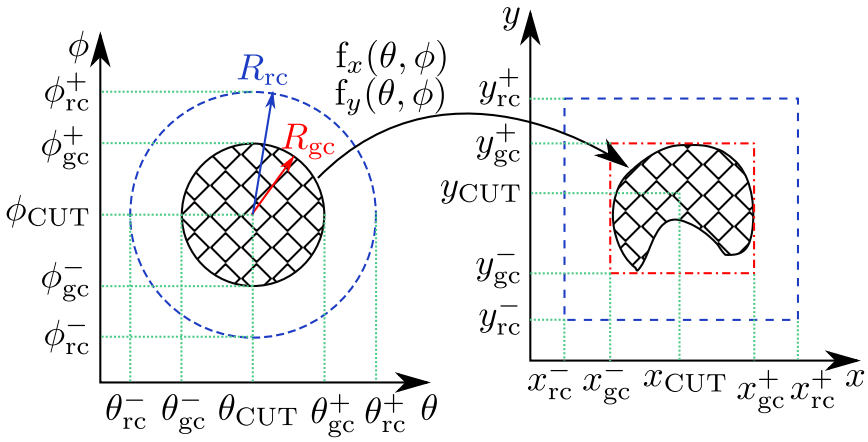

FIGURE 4.8: Transformation of the adaptive window and its parameters.

The threshold constant τ and the test statistic T_{CUT} are then calculated using these local window and guard cells size and assigned to each CUT. The adaptive CFAR threshold is then calculated in the usual way, but using these locally calculated τ and T_{CUT} parameters.

4.4.3 Direct application of CFAR in 3D space

The second approach is to directly apply a 3D CA-CFAR processor to the intensity information in the 3D point cloud data. The 3D range, bearing and elevation coordinates are first converted to x, y, z coordinates in Cartesian space. This has the advantage of not having the distortions introduced by the projections and no "maximum range" value needs to be used. On the other hand, special considerations need to be taken to apply the CFAR algorithm on the sparse 3D data, this will be described later in this section. In order to illustrate this proposed approach, Figure 4.9 shows a 3D representation of the CFAR windows operating on a particular set of points of the 3D point cloud. The sensor discretises the space into 3D voxels for each range, bearing and elevation bin. The blue point represents the point in the point cloud corresponding to the current Cell Under Test (CUT). The gray voxels represent the discretised space around the CUT point. Points in the point cloud which are closer to the CUT point than the *guard cell radius* are considered guard points and are coloured red. The red translucent sphere, with radius equal to the *guard cell radius* corresponds to the *guard window*. All points inside this sphere are guard points (red coloured). On the other hand, the green translucent sphere represents the *reference window*. All points inside this sphere and not in the *guard window* (i.e. outside the red transparent sphere) are considered reference points (green coloured), and are used in the CFAR calculations. The black points in the Figure correspond to points in the point cloud which are not taken into consideration for the current CUT point calculations, because these points are out of the guard and reference windows, i.e. too far away from the CUT point. The 3D CFAR algorithm iterates over all 3D data points in the point cloud. Each point becomes the CUT once. Then all surrounding data points which fall into the *guard window* are ignored, while the data points in the *reference window* are used to estimate the adaptive threshold. In the particular case of CA-CFAR, the reference average intensity, $T^{\text{3D}}_{\text{CUT}}$, used to estimate the threshold can be obtained by

$$T^{\text{3D}}_{\text{CUT}} = \frac{1}{V_R} \sum_{\mathbf{x}_i \in R} I_{\mathbf{x}_i} \tag{4.15}$$

where $I_{\mathbf{x}_i}$ is the intensity value corresponding to the data point $\mathbf{x}_i$, R represents the *reference window*, and V_R the number of reference points in R. Analogously, the OS-CFAR's $T^{\text{3D}}_{\text{CUT}}$ can

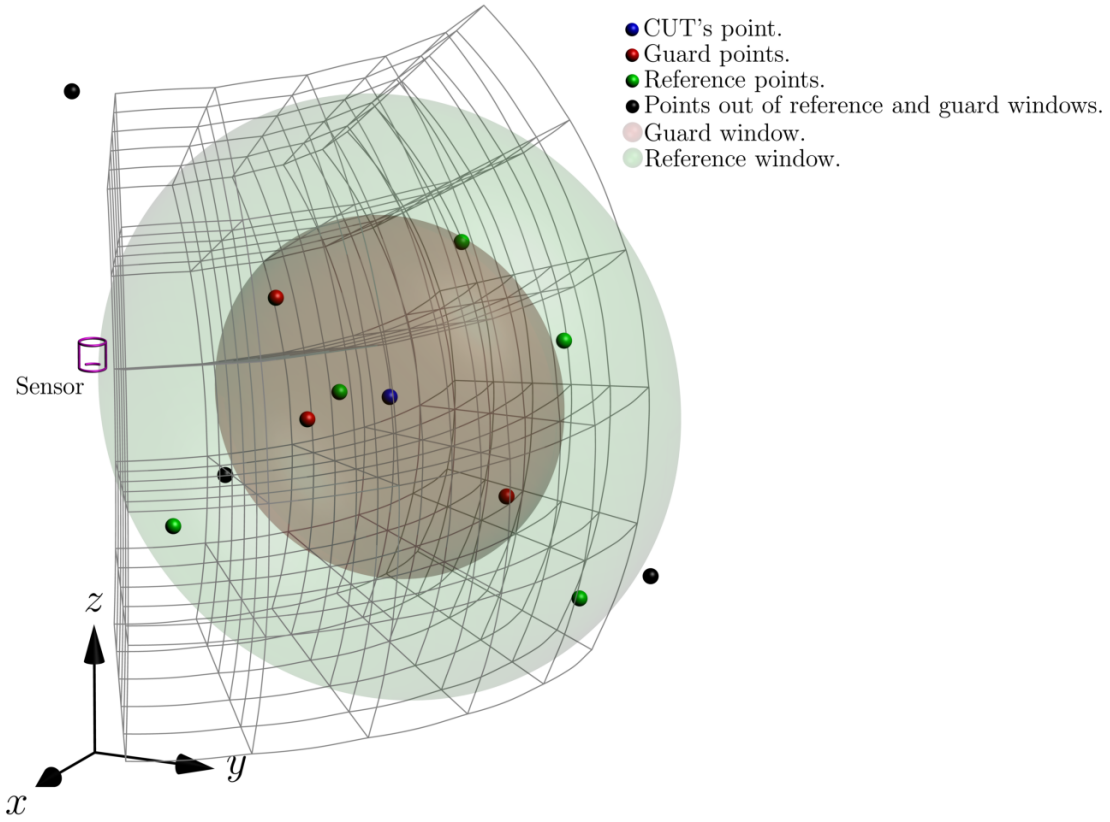

FIGURE 4.9: CFAR windows in 3D space. The point corresponding to the current Cell Under Test (CUT) is represented by a blue point in the center. The sensor discretises the space in 3D cells represented by the gray voxels (only a few voxels close to the CUT point are drawn). The red translucent sphere represents the *guard window*. All (red) points which fall into voxels, whose centre coordinates are inside this window are guard points. The green translucent sphere represents the *reference window*. All (green) points which fall into voxels, whose centre coordinates are inside this sphere but outside the *guard window* are reference points. Points outside the reference window (black coloured) are not included in the CFAR calculation for the current CUT point.

be calculated as described in Section 2.2.2. For other CFAR methods, a similar approach can be used. The CFAR methods applied directly to the 3D space avoid the distortions introduced by the 2D projections. Furthermore, the exact location of each measurement is used, while in the 2D projections, usually interpolation is used to assign the value to each pixel in the 2D image grid from the 3D data. It is expected that the targets' shape and size preservation will increase the CFAR detector's performance, because its window size parameters are related to the the size and distribution of the targets. In particular, it is expected that none, or a relatively small amount, of the reference cells should be part of the object being detected,

therefore the guard cells volume should be large enough to hold as much as possible of the object's cells detected by the sensor. Likewise the reference window should be sufficiently large to contain enough reference cells corresponding to noise and clutter, in order to obtain a valid estimate of the noise power. Nevertheless, the reference window should not be too large so as to include target cells from a neighbouring object[1]. Thus, by preserving the target's shape and size, there is a higher probability that the selected parameters will match the data as expected. Also, the higher complexity of analyzing a 3D point cloud can be partially compensated by avoiding the processing of data in the projection stage.

Unlike 1D radar data or 2D images where contiguous cells usually have contiguous indexes, 3D point clouds are often large, unstructured, lists of 3D coordinates. Therefore, an efficient method to organise and search points in a multi-dimensional space is a k-d tree [Ben75]. The nearest neighbours algorithm, which takes advantage of data structured in a k-d tree, is used to find the guard and reference cells. The implementation used in this work is based on the method described in [MM99].

Because of sparsity, there might be few or no points in the reference cells window. When the number of reference cells is small, the calculated cell average will give a less reliable estimate of the noise. Moreover, when there are no points in the reference cells window, the noise estimate is undefined. Therefore, points for which the number of reference cells is zero are discarded.

4.5 Benchmark algorithms and a comparison metric

All the CA-CFAR algorithms previously presented are compared against three state-of-the-art marker segmentation methods:

Segmentation thresholding is a fast and simple, yet widely used, segmentation method based on a global threshold, calculated based on the maximum and minimum intensity values in the data set. According to the sensor manufacturer, a constant threshold of 6 dB should be used to segment retro-reflective objects. Nevertheless, depending on the environmental conditions, the markers' measured reflectivity varies greatly, and the suggested threshold yields many missed detections. The particular implementation used here sets the threshold (β) at 45% of the full reflected intensity range in dB. In general, the obtained threshold is lower than 6 dB. This threshold has been found in

[1] As a rule of thumb, the reference cells window should be smaller than the expected minimum distance between targets, and it should contain enough noise samples. In the experiments presented in this work, reference cells window diameters 1.5 to 2.0 times larger than the corresponding guard cells window diameters showed good results.

previous experiments to yield good results with the ladar sensor used in this work. The threshold is therefore calculated as follows

$$\beta = I_{\min}^{\log} + 0.45 \cdot (I_{\max}^{\log} - I_{\min}^{\log}), \tag{4.16}$$

where $I_{\min}^{\log}$ and $I_{\max}^{\log}$ represent the minimum and maximum reflected intensity power in dB.

Canny edge detection remains one of the preferred edge detection methods due to its robustness in noisy data. An edge detector based on the Scikit-image's implementation [vdWSN+14] of the Canny [Can86] edge detector with automatic threshold selection is also used to compare the performance of the CFAR based methods.

The automatic threshold selection is performed as follows:

1. Compute the median ($\tilde{v}$) of the image's intensity values
2. Calculate the lower threshold t_l and upper threshold t_u as

$$t_l = \max(v_{\min}, \tilde{v} \cdot (1 - \sigma)) \tag{4.17}$$

$$t_u = \min(v_{\max}, \tilde{v} \cdot (1 + \sigma)), \tag{4.18}$$

where σ is a parameter describing how "tight" the thresholds should be (as a fraction between 0 and 1), and $v_{\min}, v_{\max}$ are the minimum and maximum possible intensity values.

After edges have been detected, a second algorithm is required to detect the expected features. In this implementation, connected component analysis (or labelling) is used to find pixel regions enclosed by edges. Regions which are smaller or larger than the expected marker size are discarded.

RANSAC shape detection The *random sample consensus* is an iterative algorithm to fit observed data to mathematical models. Since the introduction of the original method by Fischler and Bolles [FB81] RANSAC became a fundamental image processing tool. It is widely used to detect image structures which can be represented by a mathematical model, such as lines, circles, ellipses, and coordinate transformations. The algorithm randomly selects a small number of points in the data and estimates the model parameters fitting those points. Then, the model is tested against all the points in the data set. If a previously defined minimum number of points required to validate the model is achieved, then the model with its parameters is accepted, otherwise, a new sample of

points is used to estimate new parameters. This process is repeated until the required number of valid points is met or until a maximum number of iterations is reached.

The implementation from Scikit-image [vdWSN+14] is used with 2D projected data to find ellipses in the data, while the 3D point cloud data extension by Schnabel et al. [SWK07] is used to detect cylinders and sphere like objects in the 3D data.

Figure 4.10 shows sample outputs of the Canny, constant threshold and RANSAC algorithms for a given 2D projected ladar image, also shown in the Figure (top). Black points correspond to pixels identified as targets by each method. Sample outputs of the constant threshold and RANSAC 3D algorithm are presented in Figure 4.11, along with the input 3D point cloud. Red points represent the targets detected by the methods.

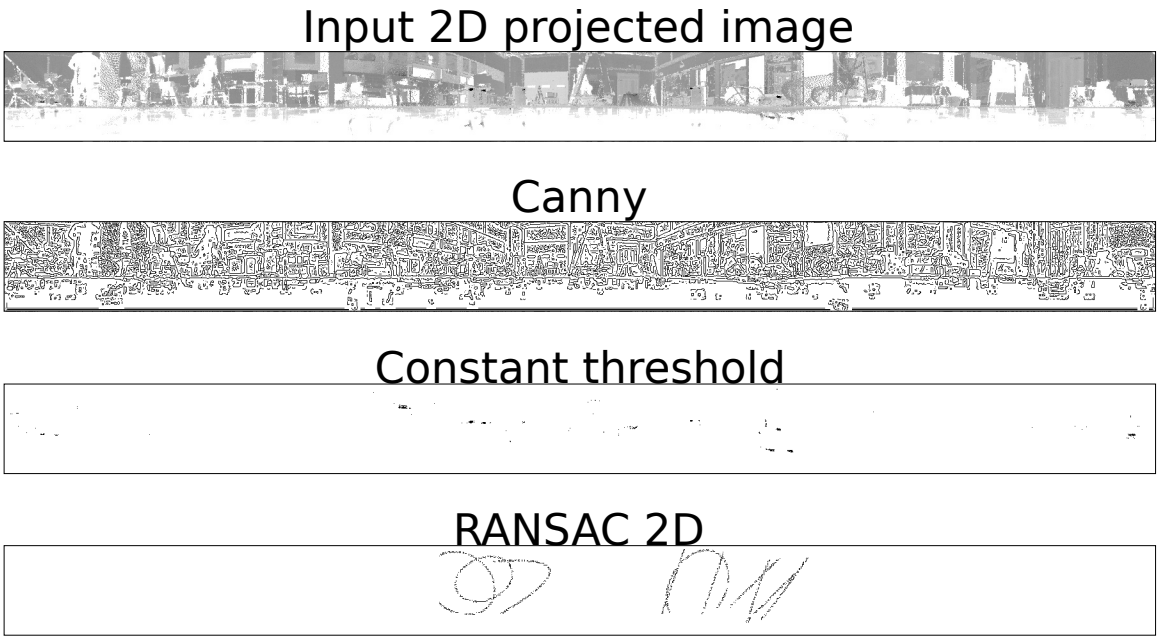

FIGURE 4.10: Sample output of the 2D benchmarking algorithm for the 2D projected input ladar image shown on top.

4.5.1 Methods' performance assessment, feature extraction and comparison metric

The performance of the methods is analysed based on multiple quantitative and qualitative indicators. Detection statistics is one of the tools used to quantify and compare the method's performance. A detection method's performance can be measured in terms of its detection

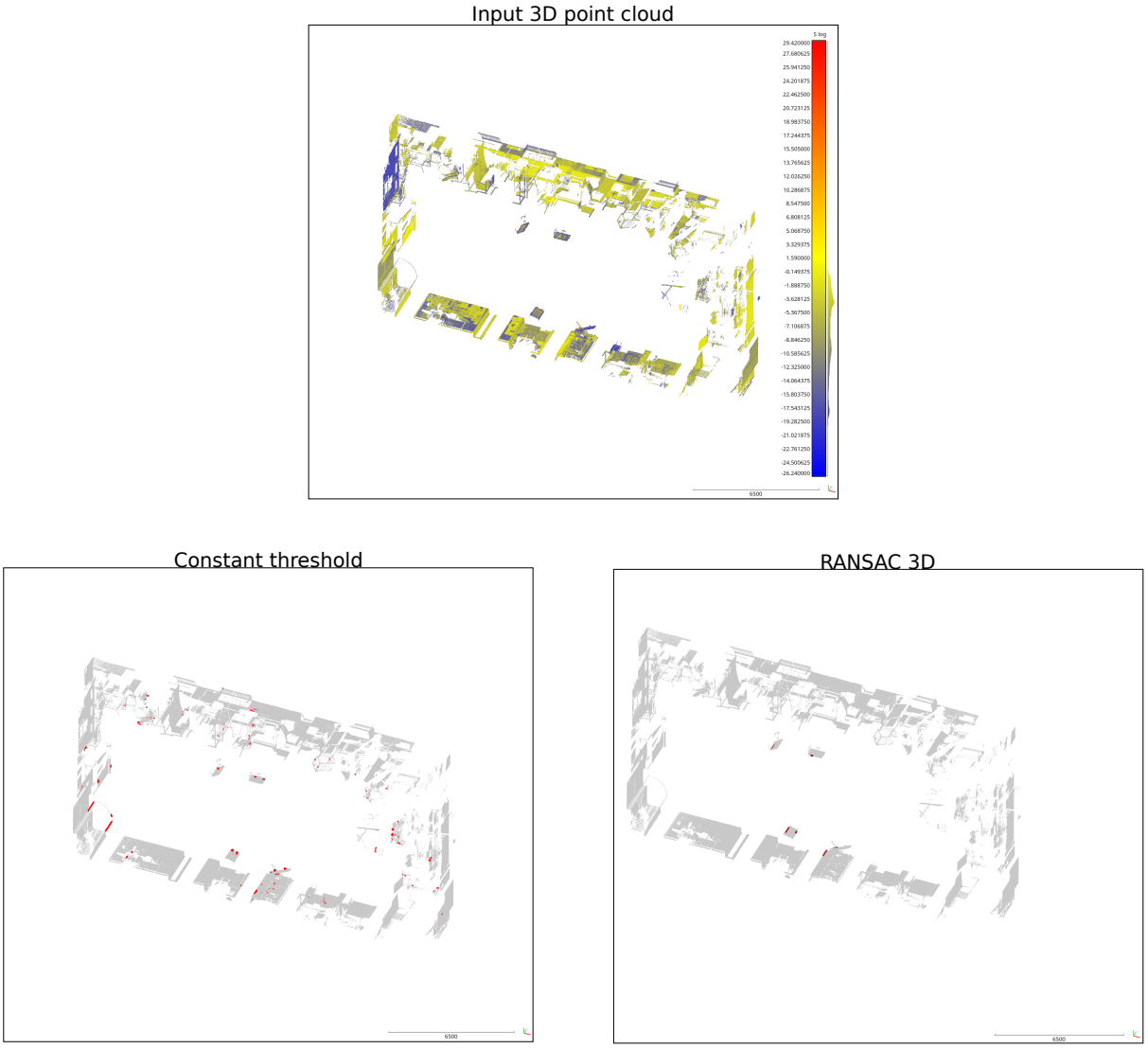

FIGURE 4.11: Sample output of the 3D benchmarking algorithm for the 3D point cloud input ladar data shown on top. Red points correspond to detections.

probability (or *true positive rate*, TPR) and its probability of false alarm (or *false positive rate*, FPR). There is usually a trade-off between TPR and FPR, which can be expressed in simple terms as follow. For any given method the probability of detection and false alarm probability are governed by some of the method's parameters. If these parameters are adjusted to increase the TPR, the FPR will also increase. Conversely, if the parameters are modified to achieve a low FPR, the probability of detection will tend to decrease.

The following definitions and equations describe the main variables and calculations in-

volved in detection statistics:

$$
\begin{aligned}
P &: \text{number of positive samples} \\
N &: \text{number of negative samples} \\
P_t &: \text{number of true positives} \\
P_f &: \text{number of false positives} \\
N_t &: \text{number of true negatives} \\
N_f &: \text{number of false negative} \\
TPR &: \text{True positive rate (probability of detection)} \\
FPR &: \text{False positive rate (probability of false alarm)}
\end{aligned}
$$

$$TPR = \frac{P_t}{P} = \frac{P_t}{P_t + N_f} \tag{4.19}$$

$$FPR = \frac{P_f}{N} = \frac{P_f}{P_f + N_t} \tag{4.20}$$

The trade-off between TPR and FPR can be studied using a *receiver operating characteristic curve*, also referred to as ROC curve. This curve is a plot of the TPR against the FPR as the discrimination threshold is varied. In a simple method, like the constant threshold algorithm, this can be achieved by changing the threshold level from the minimum value in the data (TPR and FPR are both 1) to the maximum value (TPR and FPR are both 0).

Nevertheless, most of the algorithms, both proposed and those used as benchmark, present a discrimination threshold which is controlled by multiple parameters. Some of these parameters are dependent on the context, for instance, the CFAR reference and guard cells window size is related to the distance between targets and the target size, respectively. Using the ROC curves, the sensitivity of the methods to the parameters which cannot be estimated as mentioned before can be evaluated. In this work, for the CFAR methods, the required probability of false alarm is the parameter chosen to obtain the ROC curves. The Canny automatic threshold depends on the σ parameter described above, which is used to generate the ROC curves of this method. The RANSAC algorithm has a parameter, usually referred to as *residual threshold*, which correspond to the maximum distance between a data point and the model to be classified as an inlier. This parameter has been used to produce the RANSAC's ROC curves. Finally, the single threshold level parameter is used for the segmentation thresholding method.

Although some of the methods' parameters can be estimated from context information about the data, these estimated values have only been used as starting points, and the parameter space has been explored in the vicinity to find more optimal parameters.

The CFAR algorithms, which are derived from detection theory, allow further analysis with respect to detection statistics. Particularly, one of the parameters correspond to the *desired* probability of false alarm. Therefore, it is possible to plot the actual probability of false alarm against the required FPR, which is the same parameter which was used to obtain the ROC curve. Additionally, the probability of detection can also be plotted against the same variable.

Due to the high resolution of the 3D laser scanner, each target is represented by several points in the data set. In order to obtain one feature per target, the connected component labelling [SS02] algorithm is used to group points. This algorithm is a simple and efficient method to classify regions based on adjacent pixel similarities. Its output can then be used to extract geometric characteristics of these regions. Subsequently, the center of mass is obtained for each group of points and used as the position of the target. This is carried out first for the ground truth, selecting the points corresponding to targets in the 3D point cloud data manually, and then running the connected component analysis on them. Secondly, the output of the detection algorithms are processed with the connected component analysis (CCA). Finally, an appropriate metric is applied to the extracted features to compare the different algorithms. The OSPA metric described in Section 2.4 has been used for this purpose. For all experiments the power p parameter was set to 2, while the cutoff parameter v was chosen according to the specific characteristics of each data set, as detailed in the following sections.

4.6 Simulated data

The methods were first tested with simulated environments. This allows tests under controlled conditions, particularly, the ground truth position of the markers is completely known. Also, it is possible to analyse the effects of changing specific parameters or environmental conditions. If the methods were tested directly on real data several conditions occurring simultaneously could interfere with each other, making it difficult to assess the response of the methods with regard to each of them.

Two simulated environments were used to test the methods. One consists of eight targets at the same distance from the sensor, surrounded by a noisy data background. Two versions of this data set have been used. In the first one, all the targets have different signal-to-noise ratios (SNRs), which are described in table 4.2. The second one is a "low SNR" version, with all the targets' SNRs set at 2.

Figure 4.12 depicts the scene. The dark points represent noise or clutter data surrounding the targets. The reflected intensity values of the noise background's data points were generated

TABLE 4.2: SNR values used in the simulated environment with equidistant targets

Target number	1	2	3	4	5	6	7	8
SNR	2	10	100	1000	10000	100000	1000000	10000000

using an exponential distribution. The circled points represent the eight targets equidistant from the sensor. Their reflected intensity values were generated from a Rayleigh distribution, as justified in Section 4.3.1.

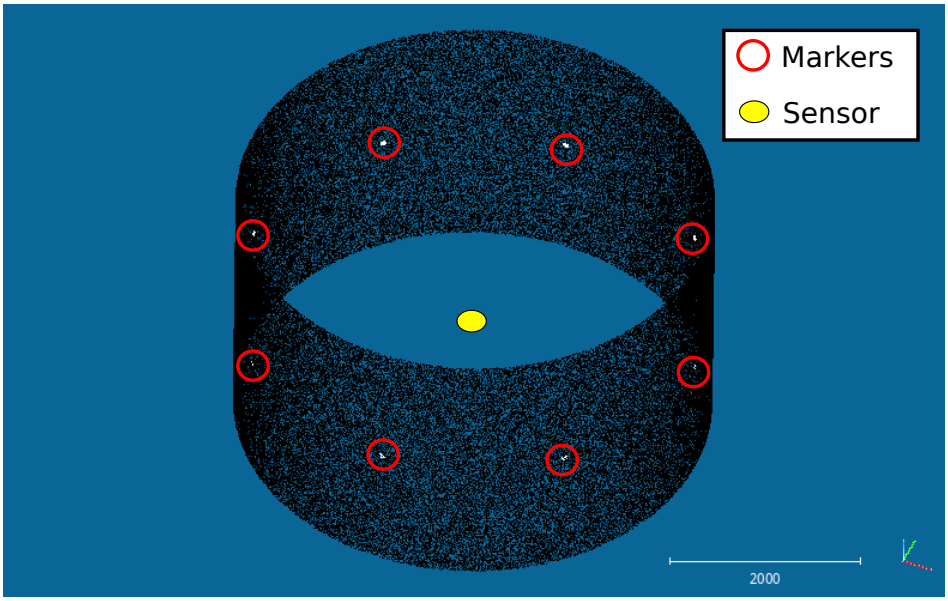

FIGURE 4.12: Simulated environment with equidistant targets to the sensor with different SNRs.

The goal of this environment is to assess the performance with respect to different SNRs, at constant range, for both target and clutter points.

The second environment consists of 4 sets of 5 targets, with all 5 targets of each set presenting the same SNR, but different SNRs for each set. The SNR for each set is listed in Table 4.3. The 5 targets of each set are at different distances, such that with respect

TABLE 4.3: SNR values used in the simulated environment with targets at different distances.

Target set	A	B	C	D
SNR	2	10	100	10000

to the sensor, each target has an apparent size[2] which doubles that of the previous closer target. Figure 4.13 shows this environment. Dark and light points represent the same as in the previously described scene. This simulated environment evaluates the performance of the detector with respect to the apparent size, at the same SNR in each target set. The standard 2D CFAR methods in particular are sensitive to the target size since it is used to define the guard cells window size parameter. This data set is therefore useful to assess the utility of the proposed 2D adaptive window size version of the CFAR methods.

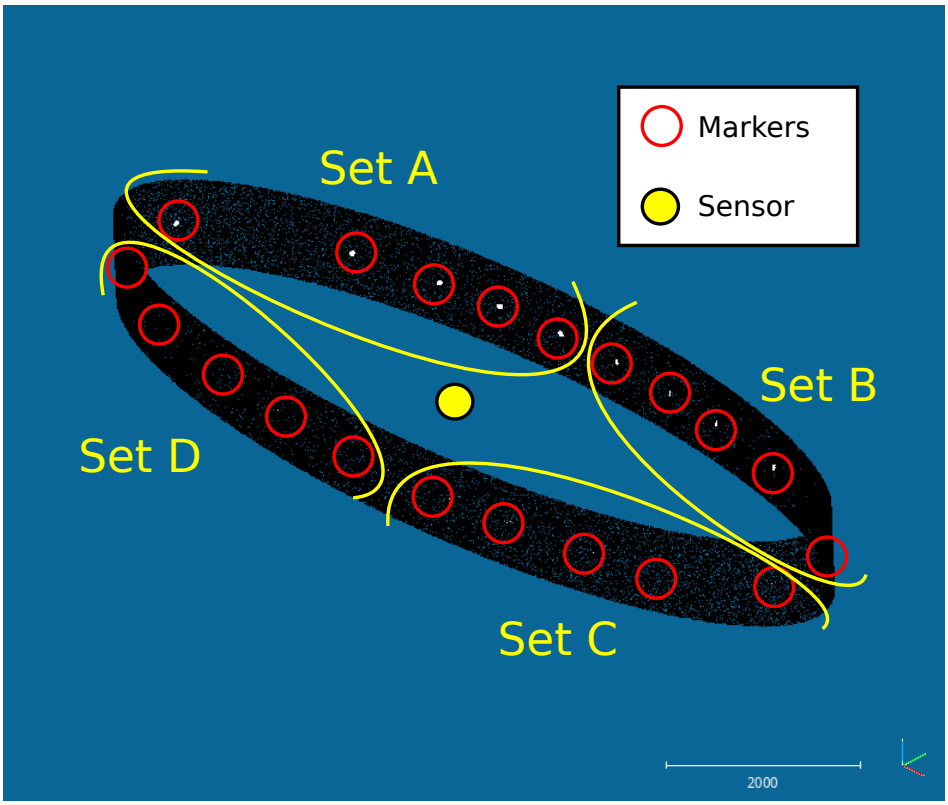

FIGURE 4.13: Simulated environment with targets at different distances from the sensor and different SNRs.

4.6.1 Equidistant target simulation results

In all the results presented in this Chapter the methods are labelled canny, cth, RANSAC, cacfar and oscfar for the Canny edge detector, segmentation thresholding, RANSAC, CA-CFAR and OS-CFAR, respectively. Additionally, the label includes the tag 2D or 3D to differentiate the respective version. CFAR 2D methods are suffixed by *std* to indicate the

[2]The apparent size is the size as seen from the sensor. Two objects of the same size will have a different apparent size if they are located at different distances from the sensor, the closer one will appear larger.

standard (fixed window size) version or *adp* to note the adaptive (adaptive window size) version.

Figure 4.14 presents the ROC curves for the 3D methods applied to the data set. The constant method 3D and the CA-CFAR 3D methods exhibit a similar performance, while the OS-CFAR 3D presents a slightly lower detection probability at low values of false positive rate (FPR) but converge with the other two for higher FPR values.

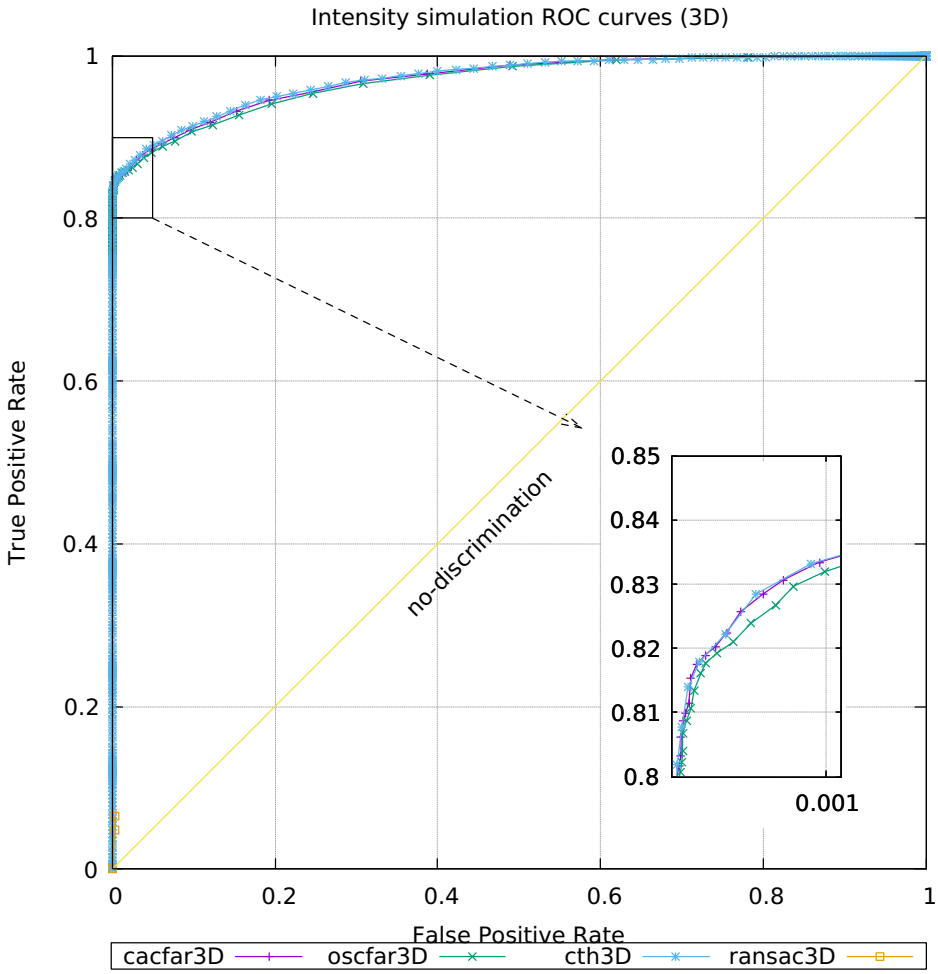

FIGURE 4.14: ROC curves of the methods applied to 3D PCD in the equidistant targets data set.

On the other hand, Figure 4.15 shows the results for the low SNR version of the data set. In this case, the constant segmentation thresholding and the CFAR methods are all detrimented due to the low signal-to-noise ratio (SNR). Nevertheless, segmentation thresholding exhibits a slightly lower performance than the other two, particularly for low FPR values. In general, in situations with high SNR, the segmentation thresholding method presents a good performance because the noise and target power distributions are well separated, thus

it is possible to find an appropriate threshold level to discriminate targets from noise. On the contrary, under low SNR conditions, CFAR methods are expected to be more robust as they use the underlying stochastic knowledge about noise and target power distributions to calculate an adaptive threshold.

However, under more realistic situations, even with high SNR levels, it is difficult to select an optimal segmentation threshold when the SNR is unknown *a priori*. CFAR methods, however, are still able to yield good results, even though there is not knowledge of the relation between target and noise power.

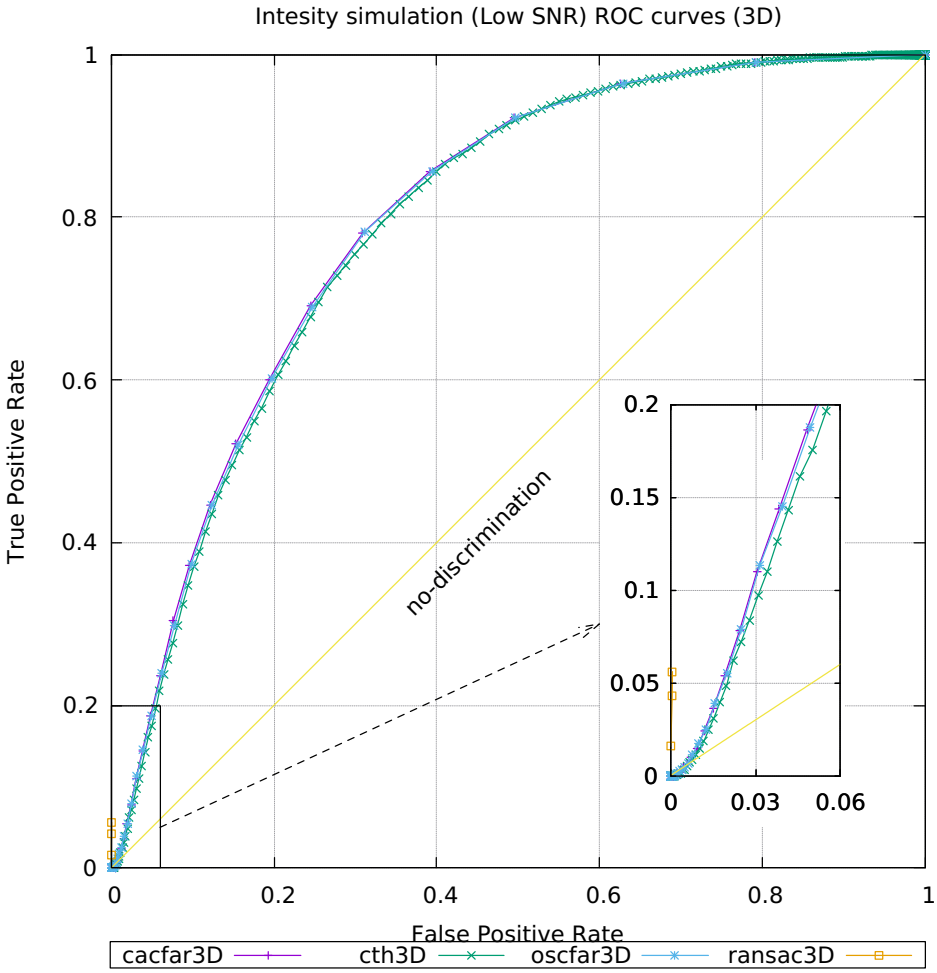

FIGURE 4.15: ROC curves of the methods applied to 3D PCD in the equidistant targets data set with low SNR.

The ROC curves of the methods applied to the high SNR version is presented in Figures 4.16 and 4.17, for equirectangular and Mercator 2D projections, respectively. It can be observed that most methods in the equirectangular projection exhibit very similar performances, while in the Mercator projection the ROC curves are more separated and towards a

lower performance level. RANSAC 2D and Canny methods both show a poor performance compared to the other methods. It is important to note that the RANSAC 2D method for ellipse detection has a very high computational complexity. For this reason the RANSAC 2D algorithm implementation was given extra information that was not available to the other methods. In particular, the RANSAC 2D method was applied to small areas in the vicinity of the ground truth targets. Nevertheless, the algorithm is unable to detect the targets due to the small number of inliers. On the other hand, the Canny edge detector is highly affected by the noise.

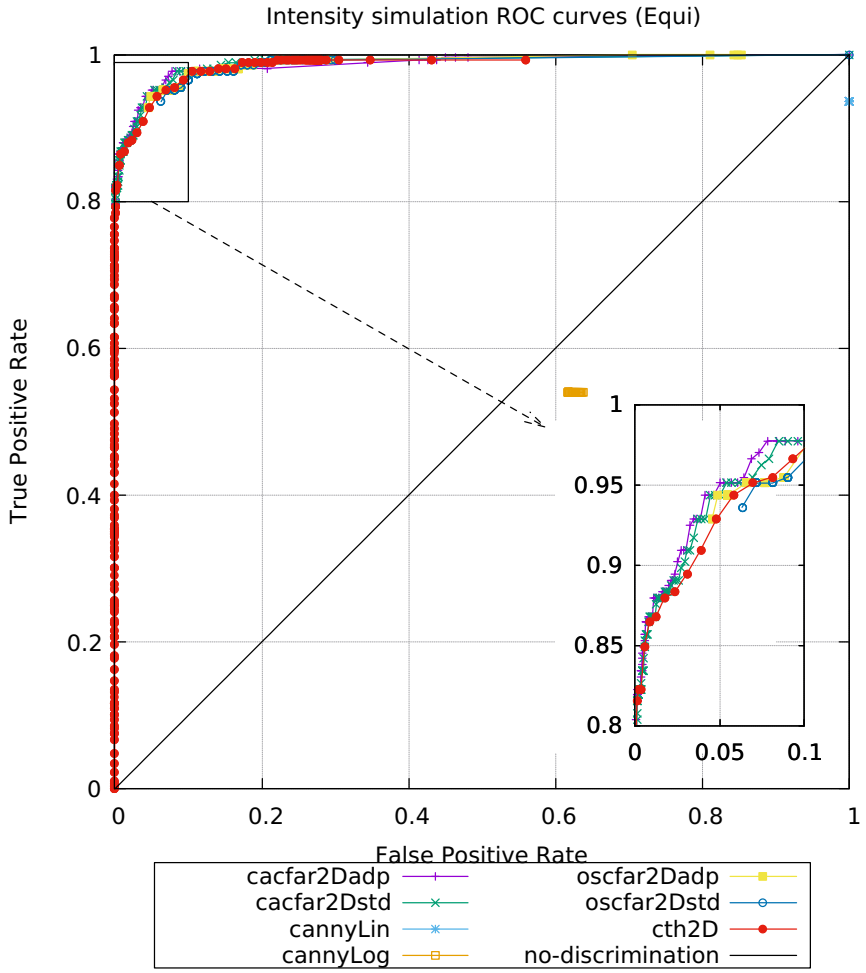

FIGURE 4.16: ROC curves of the methods applied to equirectangular 2D projected data with original SNR levels.

The 2D ROC curves for the low SNR version of the data set are shown in Figures 4.18 and 4.19, for equirectangular and Mercator 2D projections, respectively. As in the higher SNR data set, the equirectangular projection shows higher performance for all the methods than the Mercator one.

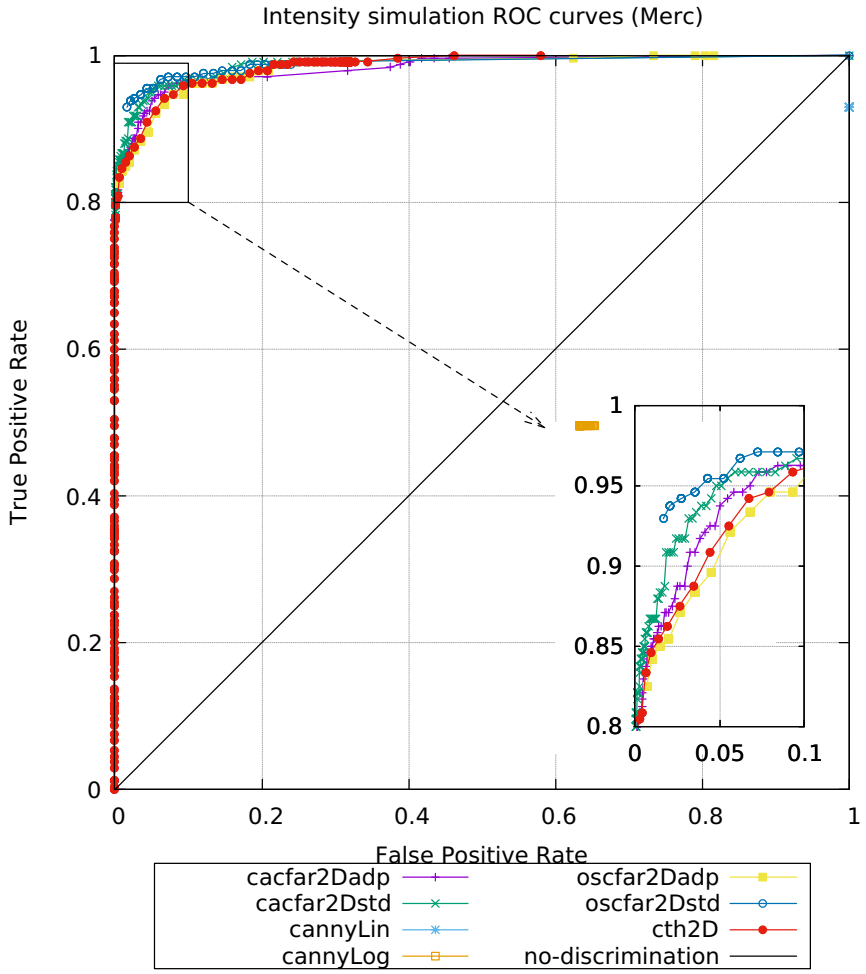

FIGURE 4.17: ROC curves of the methods applied to Mercator 2D projected data with original SNR levels.

For values of FPR < 0.2 table 4.4 ranks the methods based on the performance in terms of the ROC curves for the 2D projected data set. The curves closer to the top left corner of the plot present better performance, while those closer to the no-discrimination line are said to have a lower performance in the ROC plot.

In all cases segmentation thresholding appears in fourth or fifth rank. For both equirectangular sets (original and low SNR) only OS-CFAR 2D standard presents a lower performance than segmentation thresholding. Also, in both cases, one of the proposed adaptive versions is ranked first, either CA-CFAR 2D adaptive in the original SNR set or OS-CFAR 2D adaptive in the low SNR set. It is important to remember that the actual difference is not large, nevertheless, as it was mentioned before, it is very difficult to set the optimal threshold parameter for segmentation without knowing in advance the relation between noise and target power.

Figures 4.20, 4.21, 4.22 and 4.23 show plots of the probability of detection and actual

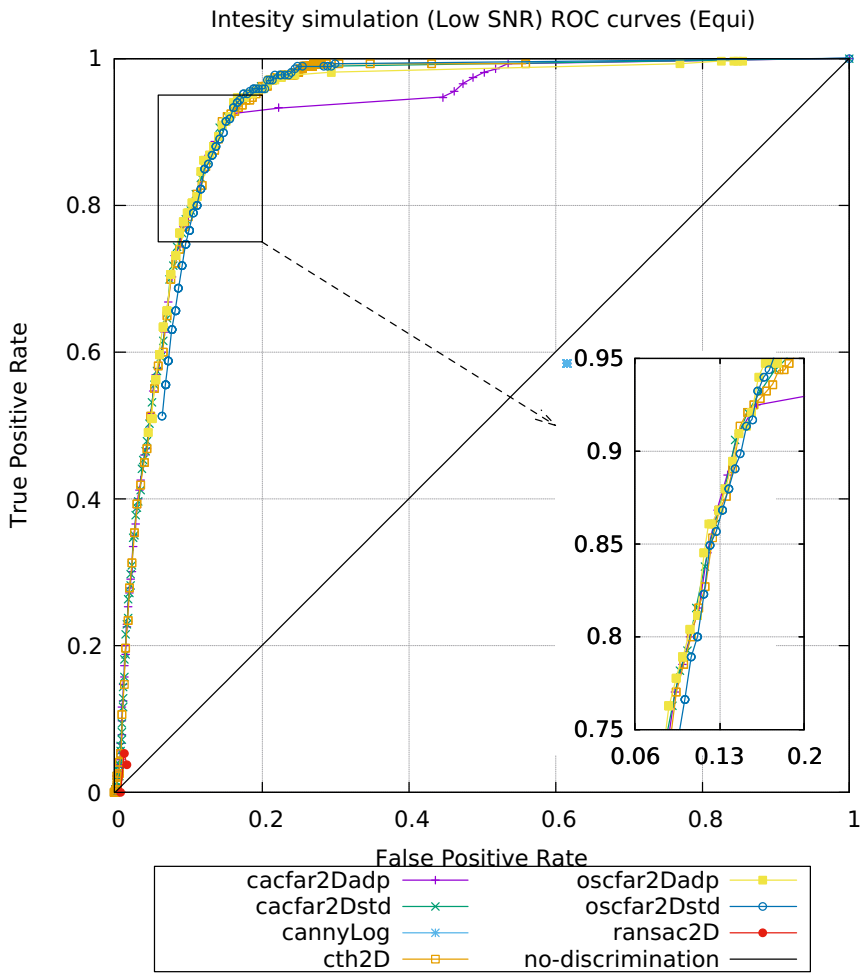

FIGURE 4.18: ROC curves of the methods applied to equirectangular 2D projected data with low SNR levels.

false alarm rate as function of the required false alarm rate, which is the main parameter that the user can set for the CFAR methods[3].

The OS-CFAR 2D (equirectangular projection), both adaptive and standard, present the higher detection probability in the original SNR set (Figure 4.20), closely followed by the OS-CFAR 2D (Mercator projection) standard. The next method in terms of probability of detection is the OS-CFAR 2D (Merc.) adaptive. Then, CA-CFAR 2D (Equi.) standard, CA-CFAR 2D (Merc.) standard and CA-CFAR 2D (Equi.) adaptive, present a comparable performance, while the CA-CFAR 2D (Merc.) adaptive shows a slightly lower detection probability. Finally, CA-CFAR 3D and OS-CFAR 3D exhibit the lowest detection probabilities of the methods. Nevertheless, all the algorithms achieve detection probabilities above 80%

[3]The other parameters are the reference and guard cells window size, which are defined based on the targets' size.

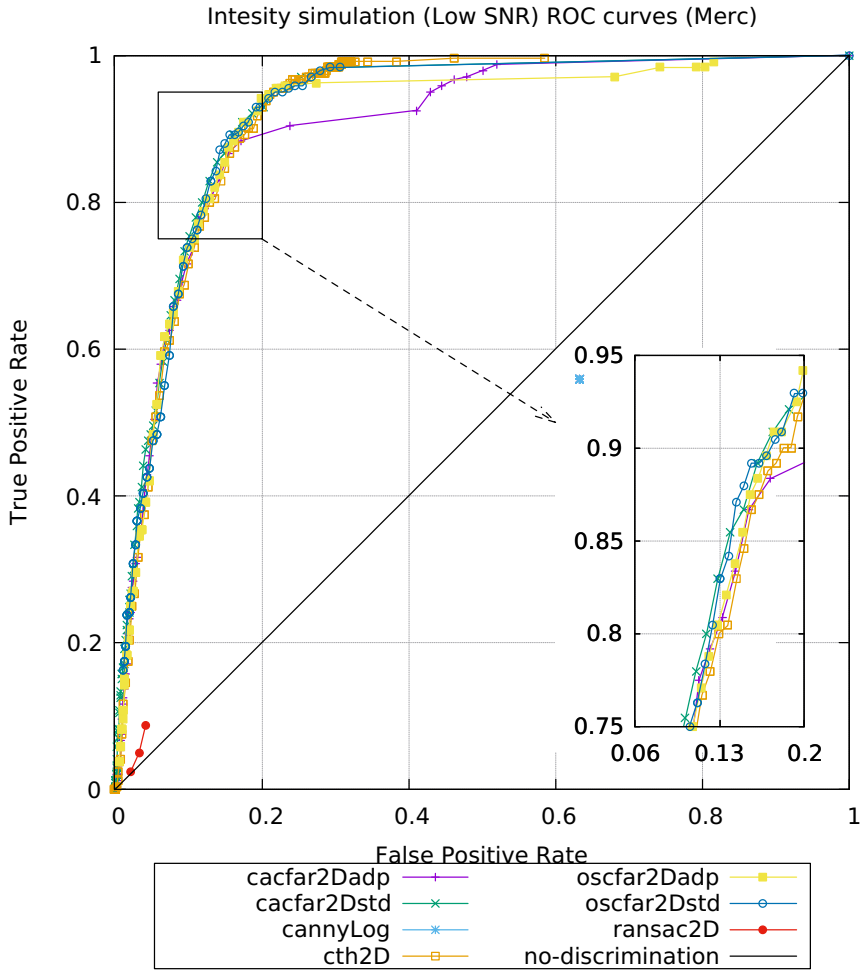

FIGURE 4.19: ROC curves of the methods applied to Mercator 2D projected data with low SNR levels.

for required false alarm rates greater than 1×10^{-4}.

On the other hand, considering the actual false alarm rate (Figure 4.21), the 3D PCD methods, both CA-CFAR and OS-CFAR, are the only ones capable of achieving a value close to the one specified as the method's parameter. A group consisting of CA-CFAR 2D (Merc.) adaptive, CA-CFAR 2D (Merc.) standard, CA-CFAR 2D (Equi.) adaptive and CA-CFAR 2D (Equi.) standard are next. Then, OS-CFAR 2D (Merc.) adaptive and OS-CFAR 2D (Merc.) standard follow, while OS-CFAR 2D (Equi.) adaptive and OS-CFAR 2D (Equi.) standard are last, in terms of the actual probability of false alarm achieved by each method. Note that the order of the methods is, with a couple of exceptions, the opposite of the one based on the probability of detection from Figure 4.20. The higher values of false alarm probability are due to the higher noise introduced by the data sparsity mentioned in section 4.2.3.

Table 4.4: Ranking of 2D methods based on ROC curves for FPR < 0.2

Rank	Original SNR (Equi)	Original SNR (Merc)	Low SNR (Equi)	Low SNR (Merc)
1	cacfar2Dadp	oscfar2Dstd	oscfar2Dadp	oscfar2Dstd
2	cacfar2Dstd	cacfar2Dstd	cacfar2Dadp	cacfar2Dstd
3	oscfar2Dadp	cacfar2Dadp	cacfar2Dstd	oscfar2Dadp
4	cth2D	cth2D	cth2D	cacfar2Dadp
5	oscfar2Dstd	oscfar2Dadp	oscfar2Dstd	cth2D
6	ransac2D	ransac2D	ransac2D	ransac2D
7	canny (log)	canny (log)	canny (log)	canny (log)

The results based on the low SNR data set (Figure 4.22) show a similar grouping of the methods in terms of performance. However, the probabilities of detection are much lower for most of the methods. Only OS-CFAR 2D (Equi.) adaptive and standard versions show a detection probability close to the original SNR data set.

Figure 4.23 presents the actual false alarm rate for the low SNR data set. As expected, the CFAR methods are capable of maintaining the false alarm rate in spite of changes in the SNR level. With small differences, the methods' performances are comparable to their performances in the original SNR data set.

The CFAR parameters used in this data set are shown in table 4.5.

Table 4.5: CFAR parameters used in the simulated environment with equidistant targets. Pixel units are labelled px, while millimetre units are labelled mm.

Method	(data type)	W_{rc}	W_{gc}
cacfar2Dstd	Equirectangular	35 px	10 px
oscfar2Dstd	Equirectangular	35 px	10 px
cacfar2Dadp	Equirectangular	370 mm	100 mm
oscfar2Dadp	Equirectangular	370 mm	100 mm
cacfar2Dstd	Mercator	35 px	10 px
oscfar2Dstd	Mercator	35 px	10 px
cacfar2Dadp	Mercator	370 mm	100 mm
oscfar2Dadp	Mercator	370 mm	100 mm
cacfar3D	3D PCD	350 mm	100 mm
oscfar3D	3D PCD	350 mm	100 mm

Figures 4.24 and 4.25 show a plot of OSPA distances for the algorithms applied to 2D projected data, while Figure 4.26 presents OSPA distances for the methods applied to 3D PCD. Figures 4.27, 4.28 and 4.29, on the other hand, present the respective plots for the low

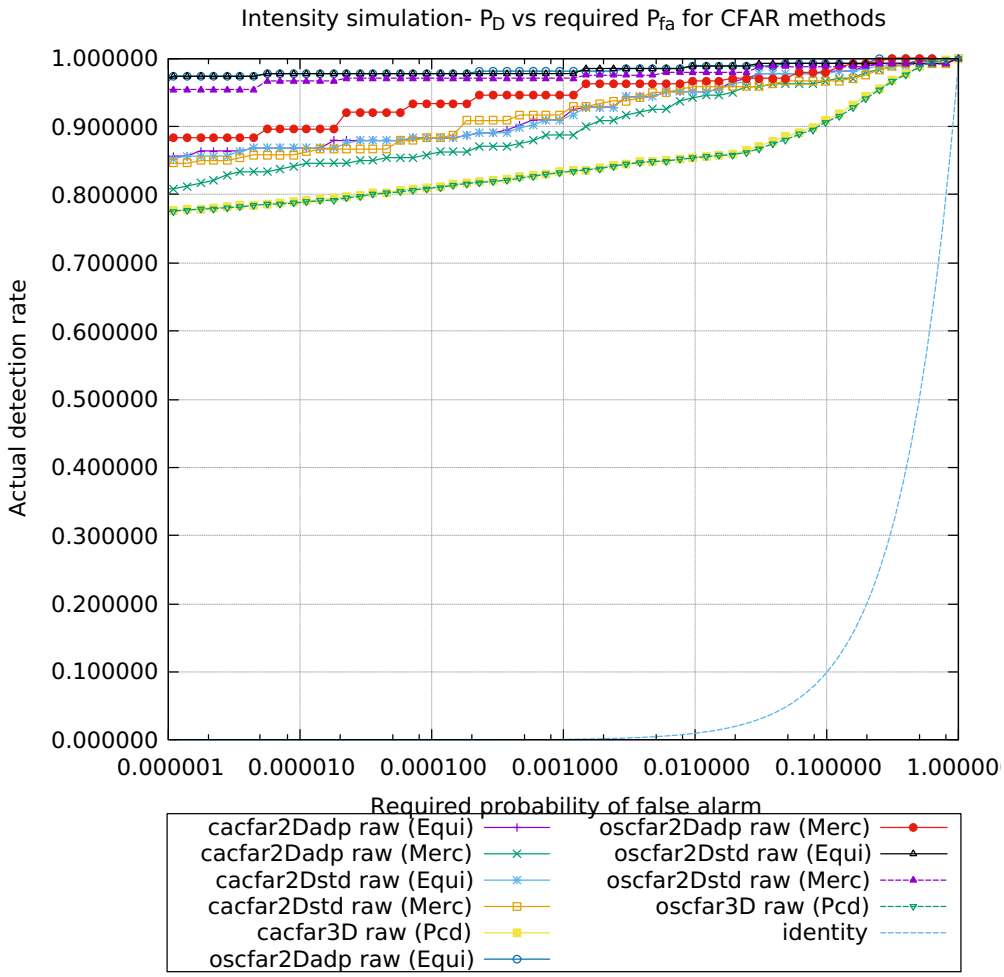

FIGURE 4.20: Probability of detection vs. required false alarm rate for original SNR data.

SNR dataset.

The OSPA cutoff parameter used with 2D projected data is different that the one used in 3D data, while the power parameter is fixed at 2 for all analyses. In the first case, the OSPA metric considers pixel distances, while in 3D it takes into account point distances in mm. In the case of the 2D projected data a cutoff parameter of 50 pixels was chosen, based on the size of the targets in the projected space and the expected separation between them. A value of 500 mm was used in the case of the 3D point cloud data, for all datasets.

In direct relation with the detection statistic results presented above, the segmentation threshold method in 2D projected data presents a lower OSPA error in the high SNR dataset. Nevertheless, for TPR values above 0.8, all methods (except Canny and RANSAC 2D) present a similar OSPA distance. Furthermore, in the Mercator projected data CA-CFAR adaptive and standard exhibit a slightly lower OSPA error.

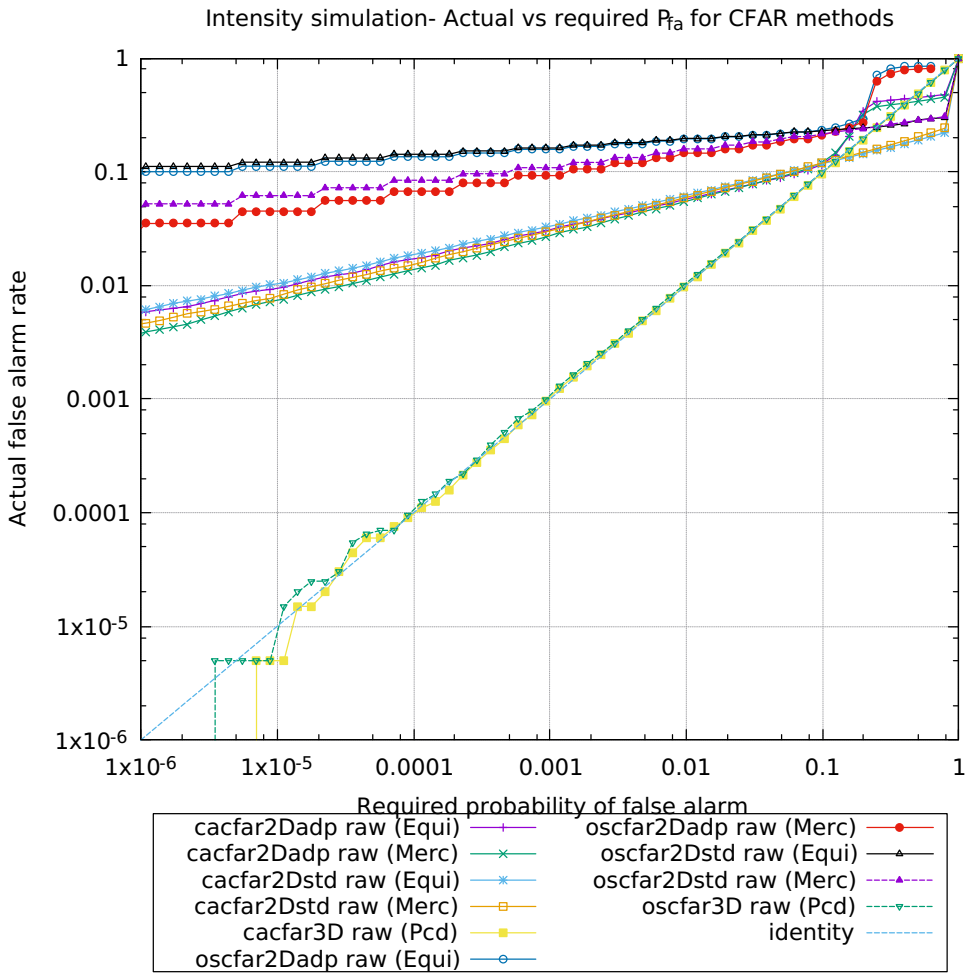

FIGURE 4.21: Probability of false alarm vs. required false alarm rate for original SNR data.

On the other hand, CFAR and constant threshold algorithms showed the same performance in the 3D point cloud data.

It should be noted that in this scenario, there were several targets with a high SNR, which should have favoured the Segmentation Thresholding method. Nevertheless, the CFAR methods outperformed the Segmentation Thresholding method and the Canny edge detector.

Also, the RANSAC algorithm, which is considered a more robust, non-deterministic method, as opposed to the other benchmarking methods, was not able yield good results. On the 2D projected data and 3D PCD, the RANSAC method follows an edge detection stage used to generate the candidate points for the ellipse model. However, the noise has two notorious effects on the RANSAC method. First, a high number of point candidates are produced. This means that more iterations are required to find a better fit. Second, the required number of valid points, can easily be reached due to the presence of false alarms.

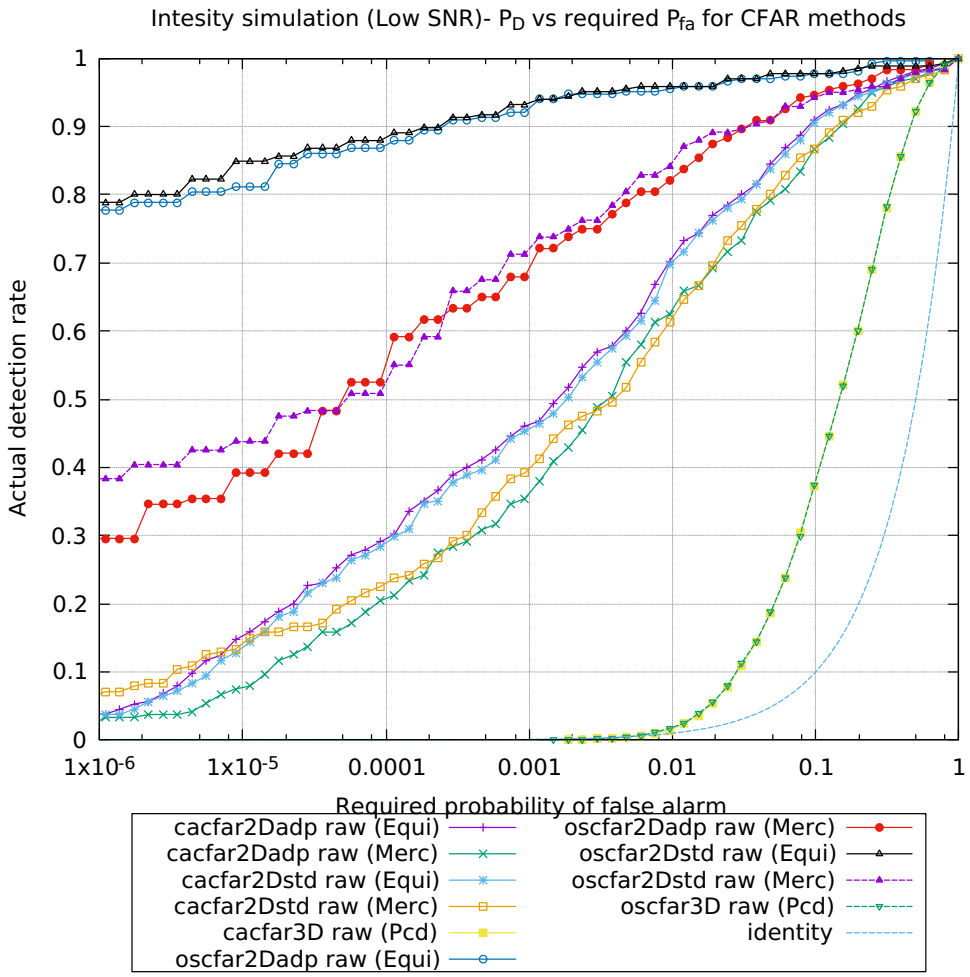

FIGURE 4.22: Probability of detection vs. required false alarm rate for low SNR data.

In the case of the low SNR data, all methods present similar performances in terms of the OSPA distance, due to the high number of false positives.

It is interesting to note that for both datasets in 3D PCD, there is a section of the OSPA distance with very low error values. This is due to the relativelly high resolution used to represent the simulated targets. Although there are is not a significative number of points detected, they are enough for the connected component analysis (CCA) method to discard the noise and keep the points of the real targets, thus yielding a low OSPA value. However, for higher TPR values, the FPR is too high and the CCA method cannot discard the noise, which increases the OSPA error.

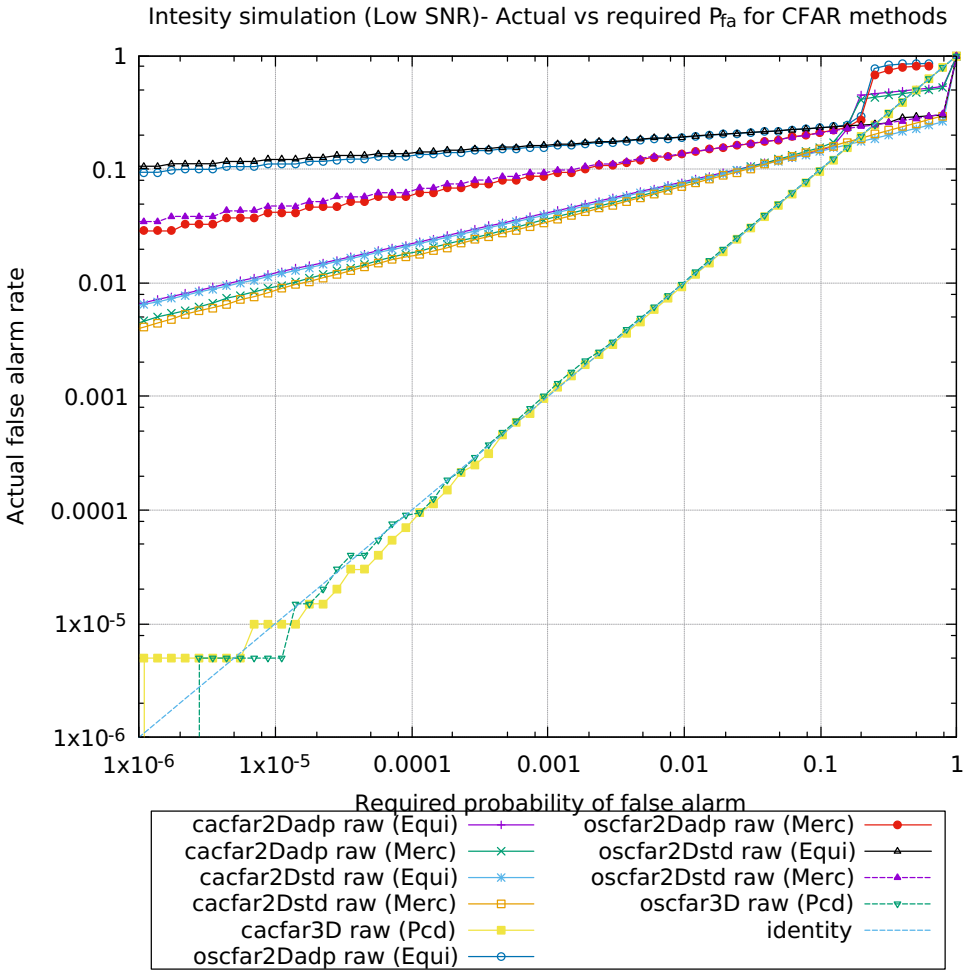

FIGURE 4.23: Probability of false alarm vs. required false alarm rate for low SNR data.

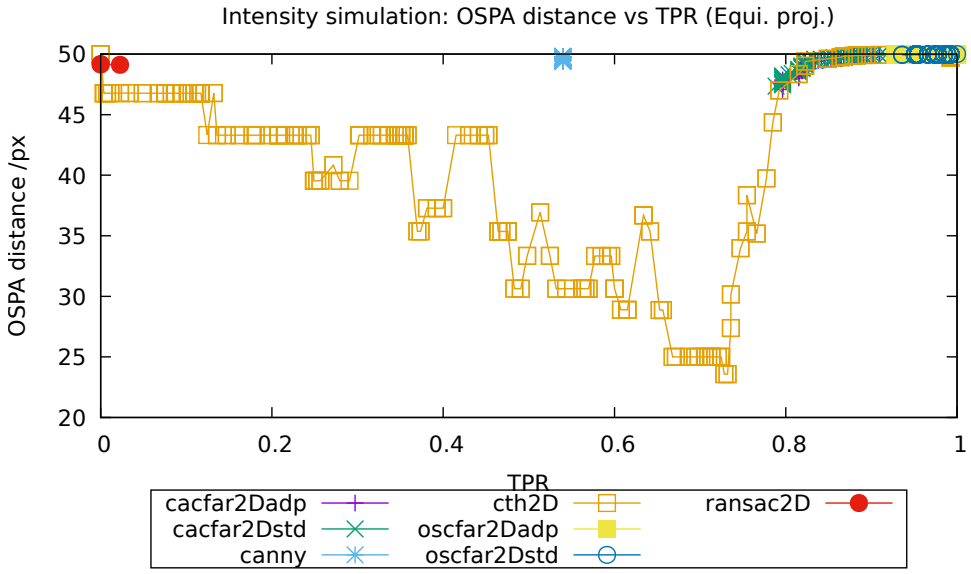

FIGURE 4.24: The OSPA metric for 2D algorithms (Equirectangular projection) in a simulated environment with equidistant targets.

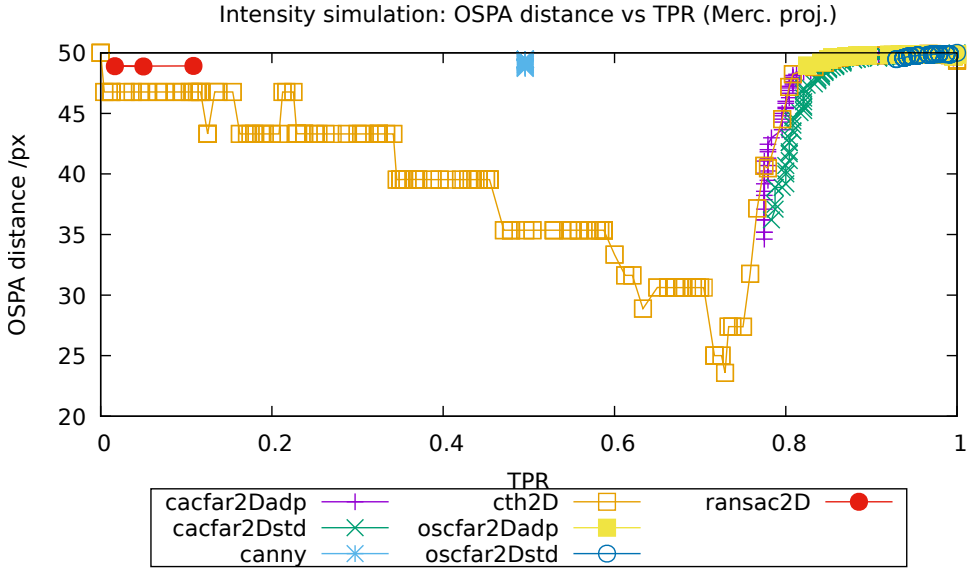

FIGURE 4.25: The OSPA metric for 2D algorithms (Mercator projection) in a simulated environment with equidistant targets.

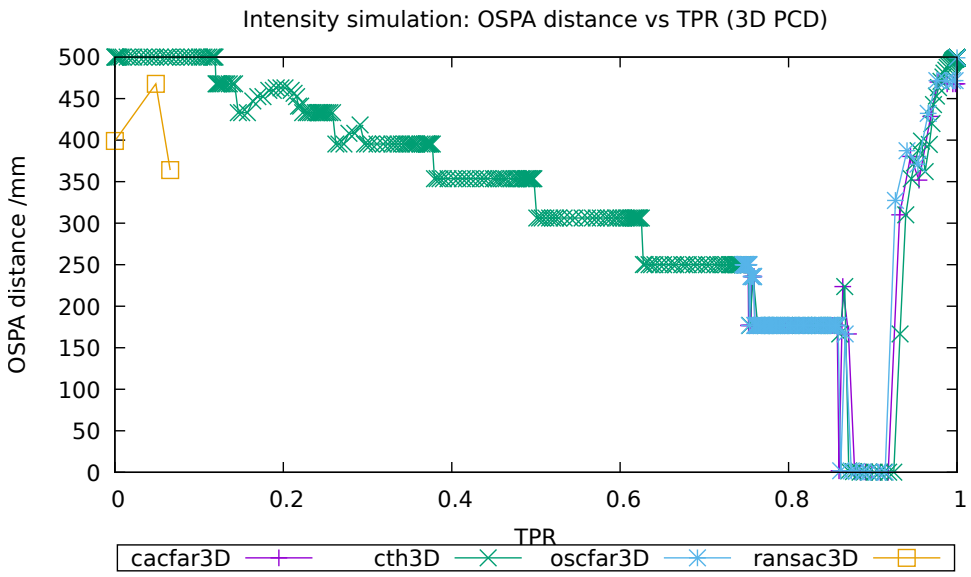

FIGURE 4.26: The OSPA metric for 3D algorithms in a simulated environment with equidistant targets.

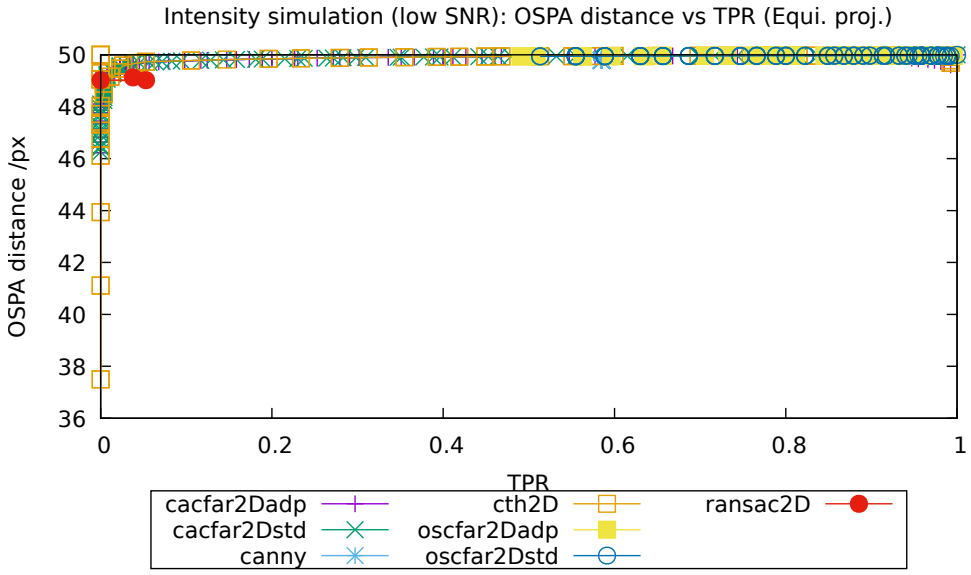

FIGURE 4.27: The OSPA metric for 2D algorithms (Equirectangular projection) in a simulated environment with equidistant targets and low SNR.

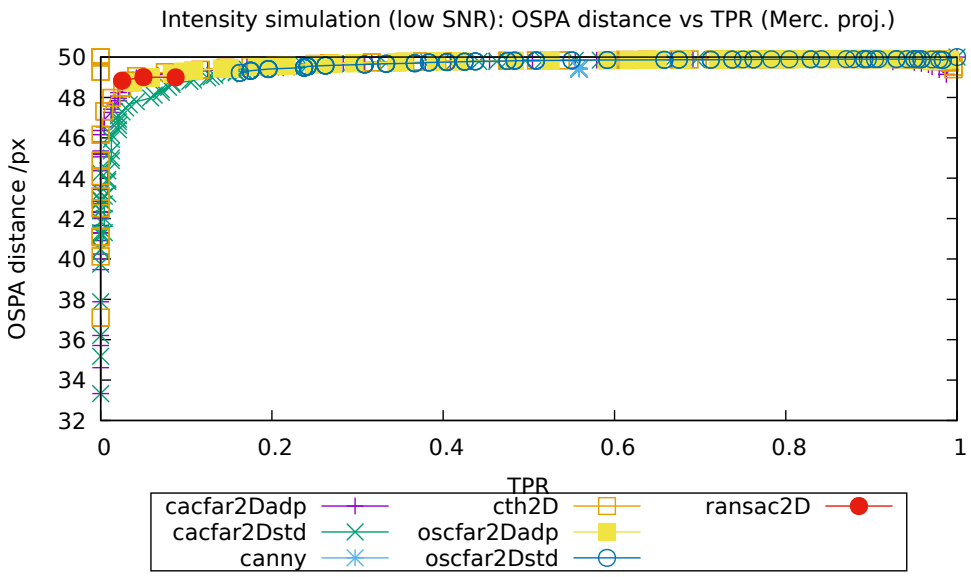

FIGURE 4.28: The OSPA metric for 2D algorithms (Mercator projection) in a simulated environment with equidistant targets and low SNR.

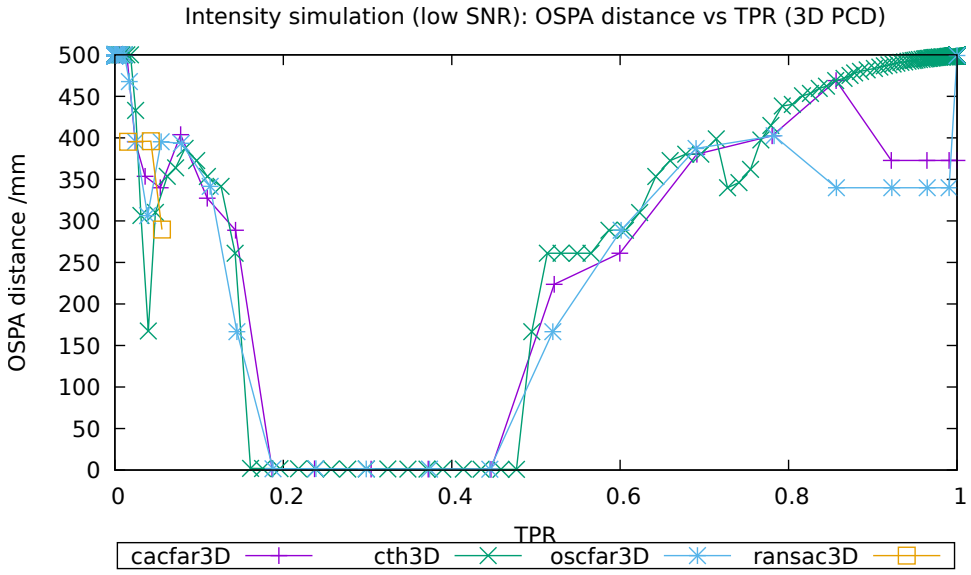

FIGURE 4.29: The OSPA metric for 3D algorithms in a simulated environment with equidistant targets and low SNR.

4.6.2 Non-equidistant target simulation results

The main goal of this data set is to evaluate the performance of the proposed adaptive window version of the CFAR methods. Thus, the 2D projected ROC curves include two versions of both OS- and CA-CFAR standard versions. In one of them the guard cells window size is chosen to tightly match the smallest apparent size of the targets, while the other one is set to match the largest apparent size of the targets. In both cases the reference cells window size is set as factor of the guard cells window size. The CFAR methods with the smallest window size are labelled SW and the ones with the largest window size are labelled LW in the ROC curves plots.

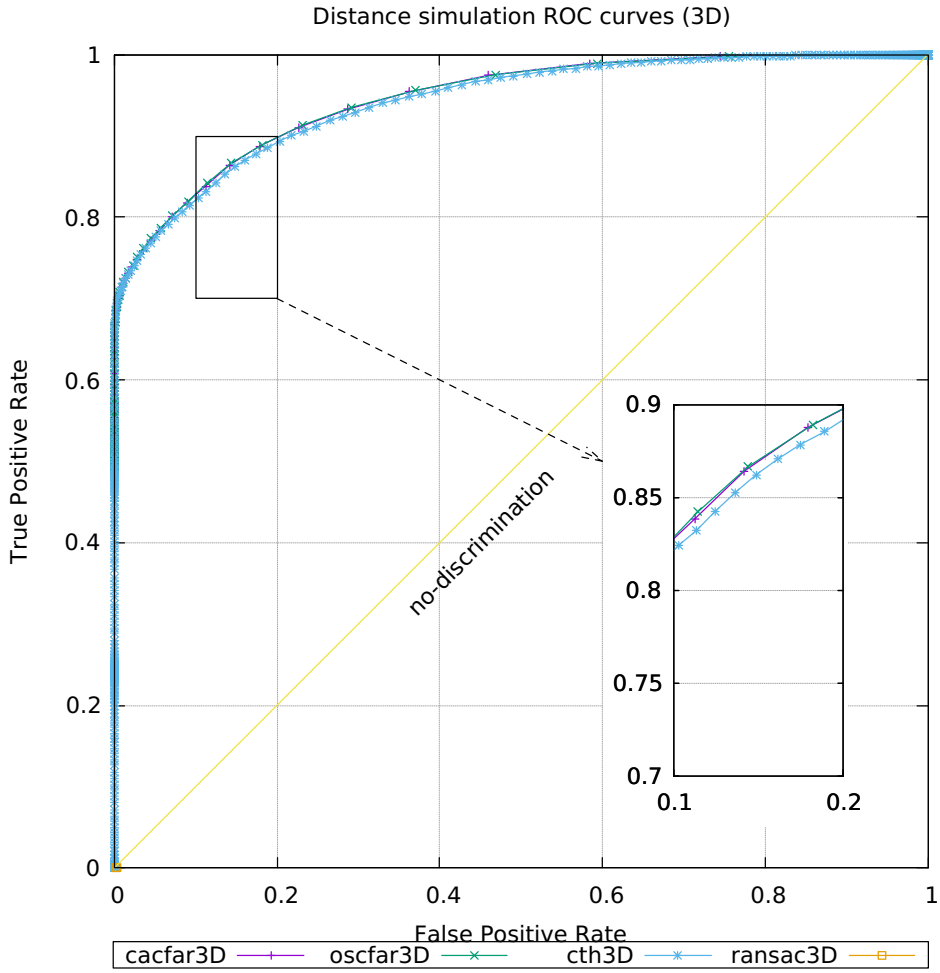

FIGURE 4.30: ROC curves of the methods applied to 3D PCD in data set with targets at different distances.

Figure 4.30 shows the ROC curves of the 3D version of the methods. In this data set, CA-CFAR 3D and OS-CFAR 3D exhibit comparable performances, while segmentation thresh-

olding presents a slightly lower performance along most of the ROC curve.

ROC curves for equirectangular and Mercator 2D projected data is presented in Figures 4.31 and 4.32, respectively.

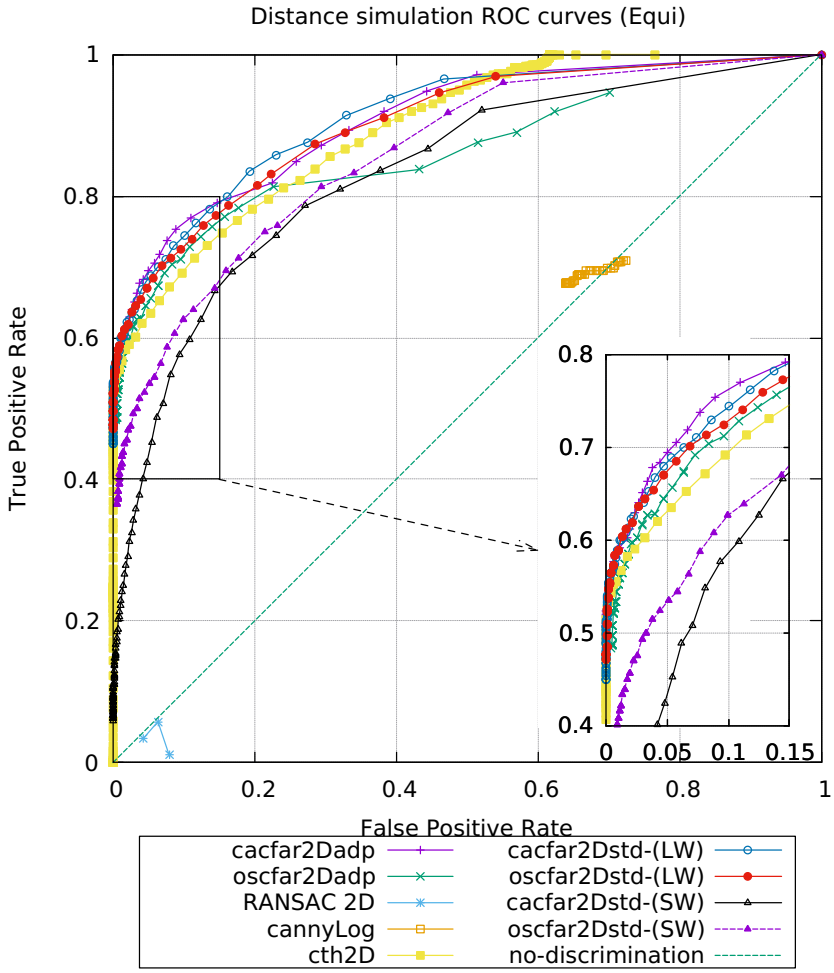

FIGURE 4.31: ROC curves of the methods applied to equirectangular 2D projected data set with targets at different distances.

In this data set both 2D projections show comparable results. Table 4.6 lists the methods ranked based on the ROC curves performances for values of FPR < 0.15.

The standard CFAR methods with a window sizes set to match the largest apparent size of targets show good results. This expected, due to the window size covering both small and large apparent size targets. On the other hand, the standard methods with small window sizes present low performances, as the large apparent size targets do not fit in the small guard cells windows. However, as in the case of the segmentation thresholding parameter, it is not possible to have a good estimate of the apparent size without knowing the relative position of the (closest) targets to the sensor.

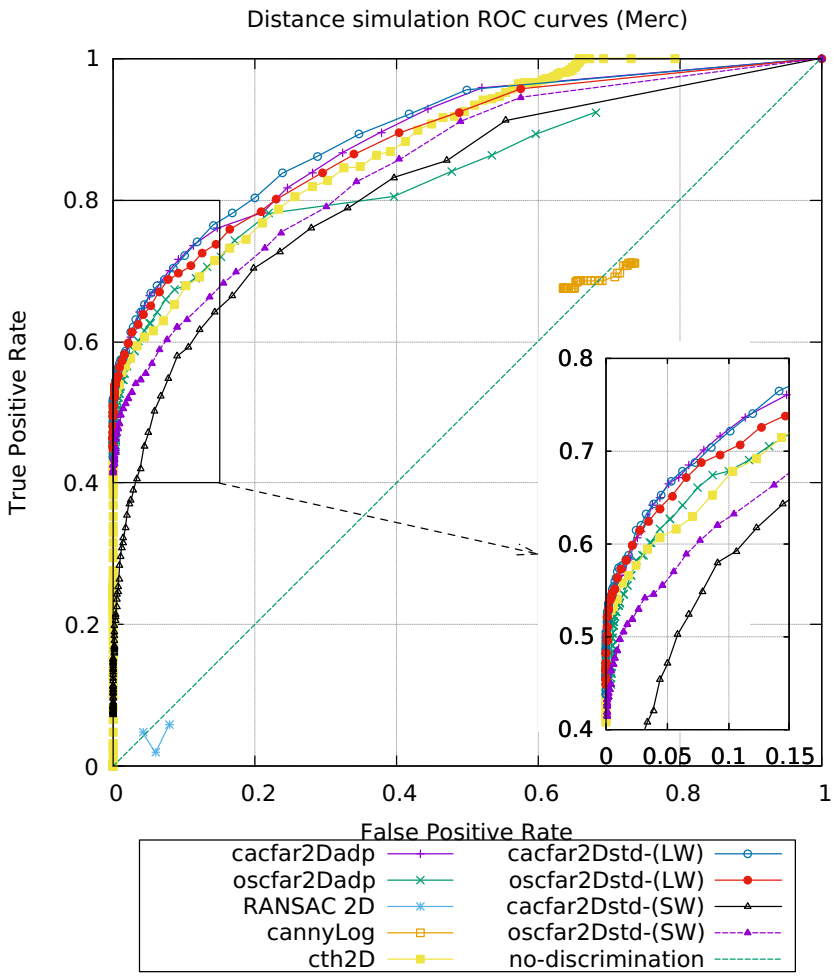

FIGURE 4.32: ROC curves of the methods applied to Mercator 2D projected data set with targets at different distances.

Figure 4.33 plots CFAR methods' detection probability versus required probability of false alarm. Likewise, Figure 4.34 plots CFAR methods' actual false alarm rate versus required probability of false alarm.

Table 4.7 ranks the methods with respect to the probability of detection and actual probability of false alarm.

It can be observed that in this data set both 3D versions of the CFAR methods present the top performance with respect to both probability of detection and actual false alarm rate. Also, the CA-CFAR 2D adaptive version (in both projections) presents a relatively good performance in terms of the actual false alarm rate, while the OS-CFAR adaptive version exhibits a low performance level. In general, the standard methods with small window sizes present low performance level, while the large window methods show mixed results. Overall, methods applied to equirectangular projected data presented better probabilities of

TABLE 4.6: Ranking of 2D methods based on ROC curves for FPR < 0.15

Rank	Equirectangular	Mercator
1	cacfar2Dadp	cacfar2Dadp
2	cacfar2Dstd-(LW)	cacfar2Dstd-(LW)
3	oscfar2Dstd-(LW)	oscfar2Dstd-(LW)
4	oscfar2Dadp	oscfar2Dadp
5	cth2D	cth2D
6	oscfar2Dstd-(SW)	oscfar2Dstd-(SW)
7	cacfar2Dstd-(SW)	cacfar2Dstd-(SW)
8	cannyLog	cannyLog
9	RANSAC 2D	RANSAC 2D

TABLE 4.7: Ranking of CFAR methods based on probability of detection and actual false alarm rate.

Rank	Detection rate	Actual false alarm rate
1	cacfar3D (Pcd)	oscfar3D (Pcd)
2	oscfar3D (Pcd)	cacfar3D (Pcd)
3	oscfar2Dadp (Equi)	cacfar2Dadp (Merc)
4	oscfar2Dstd-(LW) (Equi)	cacfar2Dstd-(LW) (Merc)
5	cacfar2Dstd-(LW) (Equi)	cacfar2Dadp (Equi)
6	cacfar2Dadp (Equi)	cacfar2Dstd-(LW) (Equi)
7	oscfar2Dstd-(LW) (Merc)	oscfar2Dstd-(LW) (Merc)
8	cacfar2Dstd-(LW) (Merc)	cacfar2Dstd-(SW) (Merc)
9	oscfar2Dadp (Merc)	cacfar2Dstd-(SW) (Equi)
10	cacfar2Dadp (Merc)	oscfar2Dstd-(LW) (Equi)
11	oscfar2Dstd-(SW) (Merc)	oscfar2Dstd-(SW) (Merc)
12	oscfar2Dstd-(SW) (Equi)	oscfar2Dadp (Merc)
13	cacfar2Dstd-(SW) (Merc)	oscfar2Dadp (Equi)
14	cacfar2Dstd-(SW) (Equi)	oscfar2Dstd-(SW) (Equi)

detection that those applied to Mercator projected data. On the other hand, it is not possible to conclude if the type of projection has some effect on the actual false alarm rate. Note that 25% of the targets in this data set present a low SNR level, which affects the detection probabilities.

In this data set the CFAR parameters used are the same as with equidistant targets.

Figures 4.35, 4.36 and 4.37, present the OSPA distances for the methods applied to 2D data and 3D PCD, respectively.

The OSPA metric results coincide with the detection statistic results. For 2D projected data, CFAR standard methods with small window (both CA- and OS-) and OS-CFAR adap-

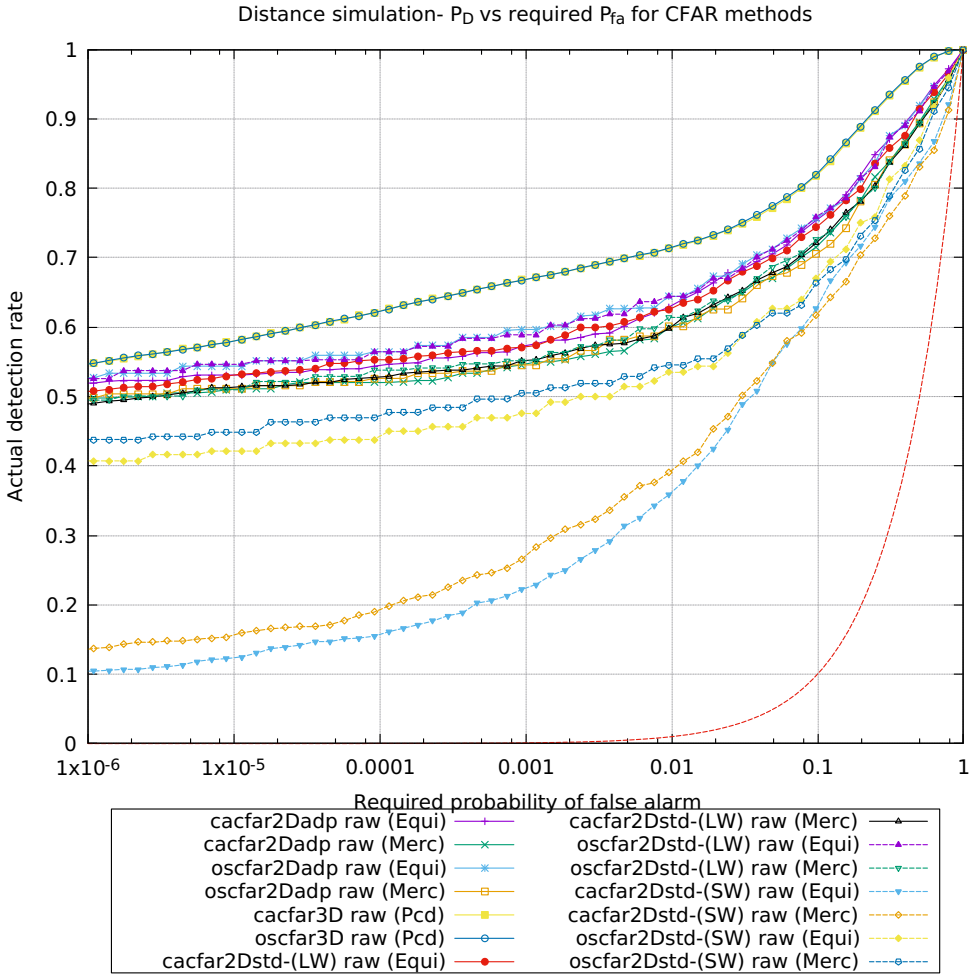

FIGURE 4.33: Probability of detection vs. required false alarm rate in data set with targets at different distances.

tive methods present higher OSPA errors because of their higher false alarm rates.

On the other hand, CA-CFAR adaptive version and CA-CFAR standard with large window show the lowest OSPA errors around a TPR value of 0.5 in equirectangular projected data. OS-CFAR standard with large window and segmentation threshold yield a slightly higher OSPA error. Nevertheless, in Mercator projected data, for TPR > 0.5, these methods present similar OSPA errors. Also, CFAR and segmentation threshold exhibit comparable OSPA performances in 3D PCD.

As mentioned earlier in this section, the goal of this simulated environment is to test the advantages of the adaptive 2D CFAR methods and the 3D version over the 2D standard versions. The 2D adaptive and the 3D versions take into account the real expected target size, while the standard version only considers one target size represented as a pixel area. In scenarios where the targets are at different distances from the the sensor, the apparent size

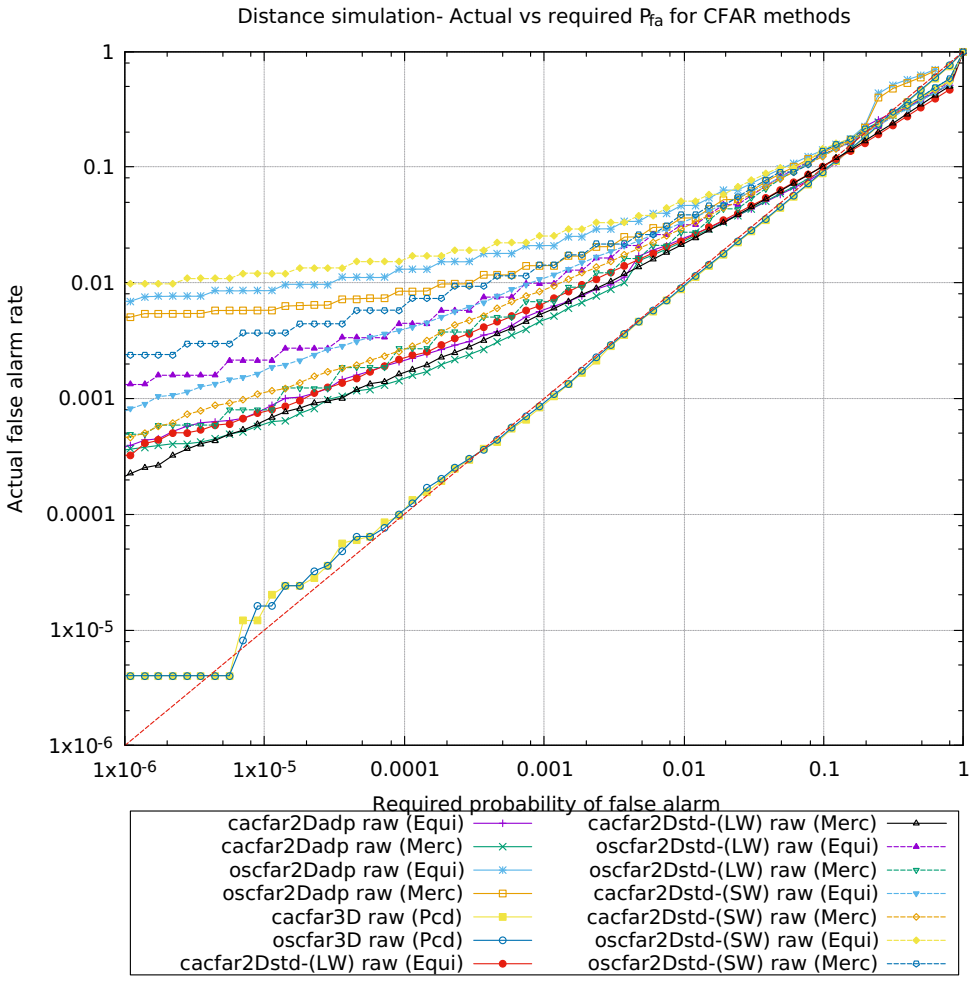

FIGURE 4.34: Probability of false alarm vs. required false alarm rate in data set with targets at different distances.

and thus the pixel area occupied by the targets vary.

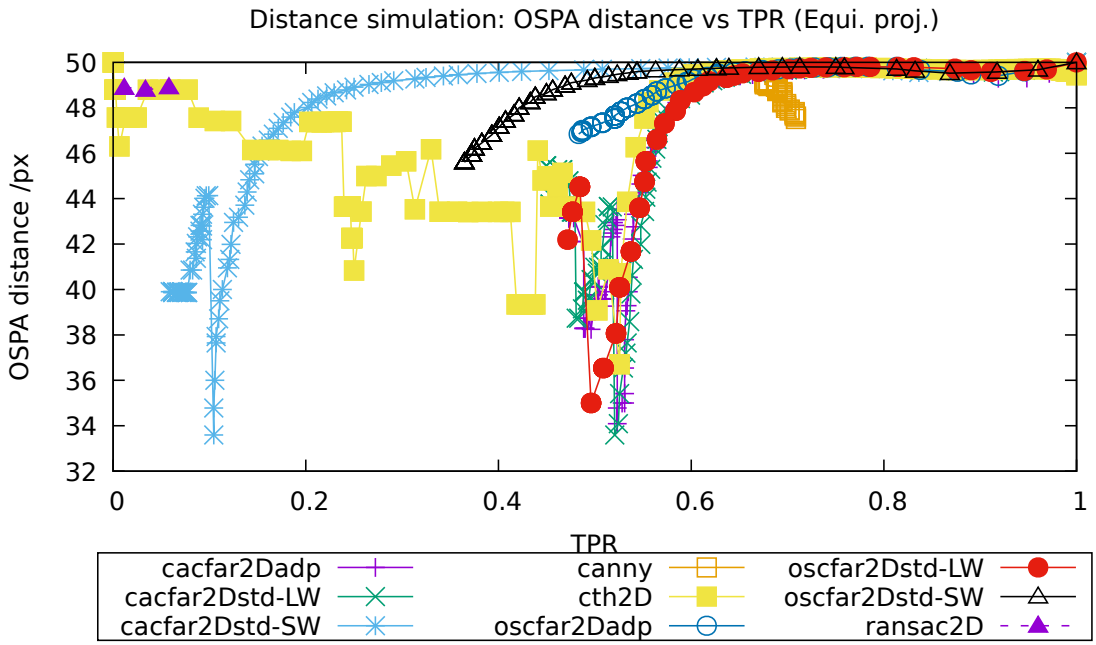

FIGURE 4.35: The OSPA metric for 2D algorithms in a simulated environment with non-equidistant targets (Equirectangular projection).

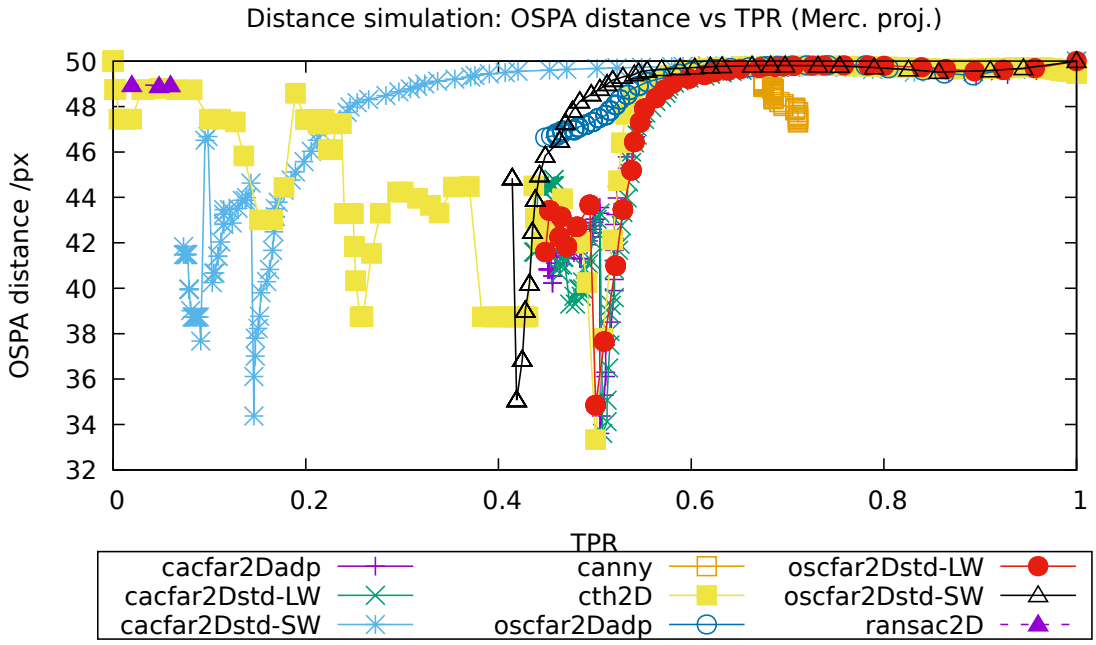

FIGURE 4.36: The OSPA metric for 2D algorithms in a simulated environment with non-equidistant targets (Mercator projection).

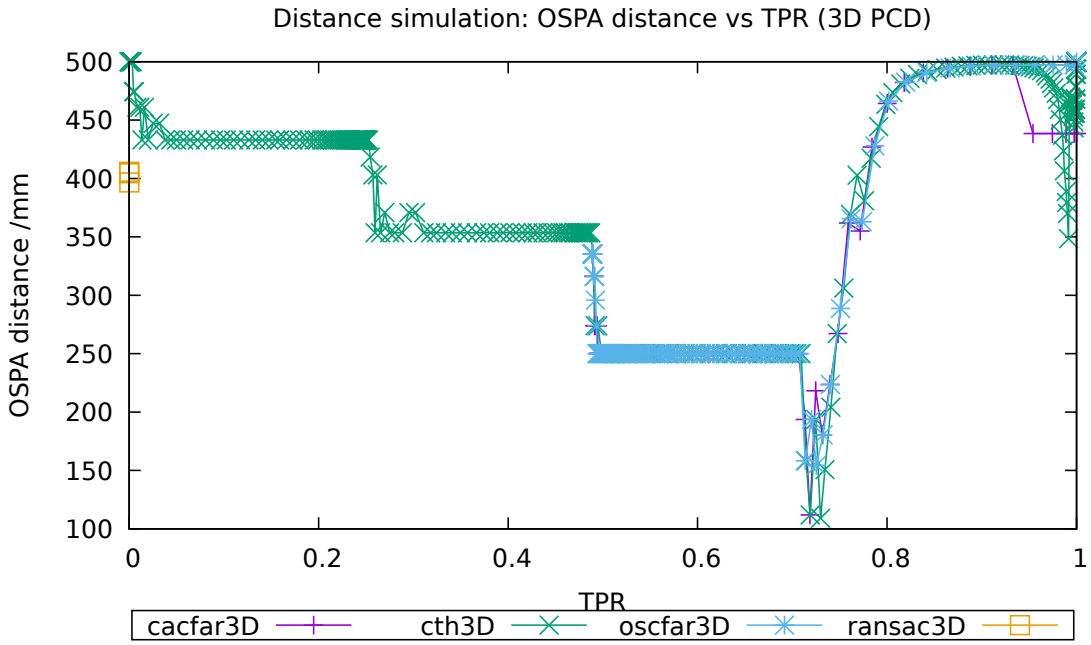

FIGURE 4.37: The OSPA metric for 3D algorithms in a simulated environment with non-equidistant targets.

4.7 Indoor laboratory environment test

The previously described methods were tested with 3D ladar data gathered in an indoor laboratory environment. Reflective markers (seven) were located in different places inside a laboratory. Figure 4.38 presents the 3D Point Cloud corresponding to the scene. Darker points represent a higher reflected intensity on a logarithmic scale.

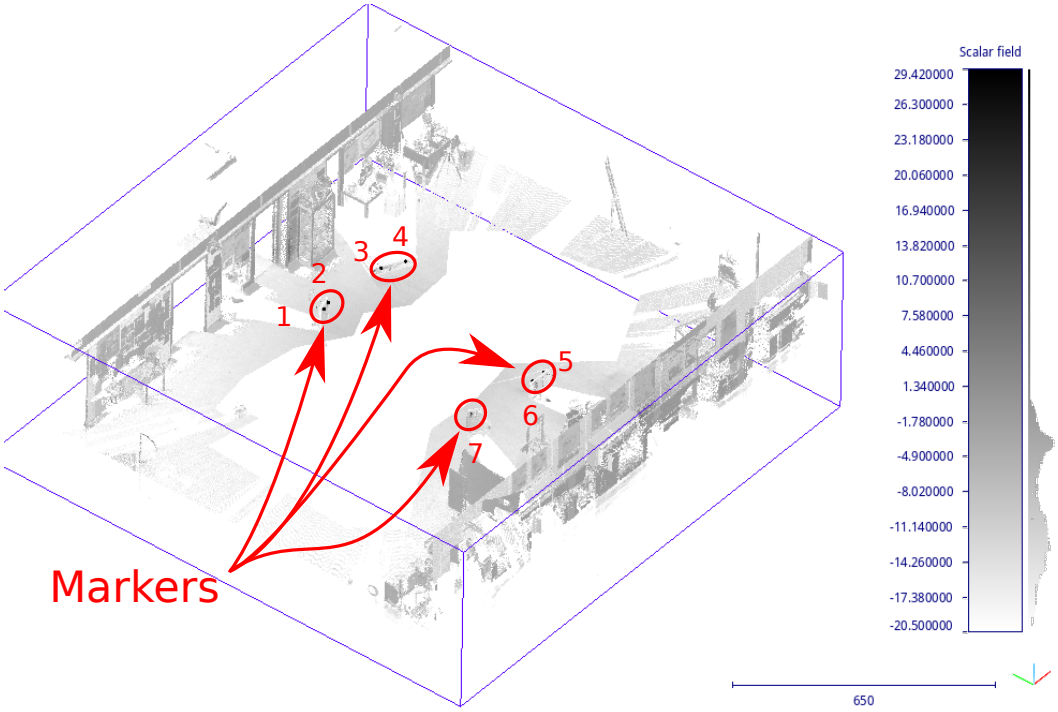

FIGURE 4.38: Indoor laboratory environment. Darker points show a higher reflected logarithmic intensity.

The reflective markers were identified manually in the 3D Point Cloud and in the 2D projected images (see Figures 4.39 and 4.40) and corresponding points and pixels were labelled to be used as ground truth. The CFAR parameters in this experiment were the same as in the two simulation experiments because the targets were of similar size in the 3D and 2D projected data.

FIGURE 4.39: Mercator projected image of indoor laboratory environment.

All experimental data sets (indoor laboratory and both section of an underground mine) present well separated distributions of target and noise power. One problem when using

FIGURE 4.40: Equirectangular projected image of indoor laboratory environment.

retro-reflective target in ladar data capture is the presence of clutter in the scene with high reflectivity which could be identified as markers. The experimental data sets do not include enough clutter to evaluate its effect on the target detection methods. Therefore, for each data set, a second copy has been created. In these secondary data sets a small rectangular area of non-target data points have been modified to artificially represent a patch of high reflectivity. This artificial clutter area has reflectivity values in the same range as the actual targets.

4.7.1 Indoor laboratory environment results

The ROC curves of the methods applied to 3D PCD for the original data set and the alternate data set are presented in Figures 4.41 and 4.42, respectively.

Due to the low SNR and the separation between noise and target power, all 3D methods show a very high performance. Both segmentation thresholding and OS-CFAR 3D achieve 100% detection rate with very small values of FPR. The CA-CFAR method, requires a slightly higher FPR to achieve 100% detection rate.

When the small patch of artificial clutter is included the segmentation thresholding is affected, increasing its false alarm rate, as expected. On the other hand, OS-CFAR 3D presents the same performance, while CA-CFAR 3D improves with a small reduction in its false alarm rate, due to some points, where the patch was introduced, being correctly detected as clutter in this scenario.

Figures 4.43 and 4.44 present ROC curves obtained from 2D data, equirectangular and Mercator projections, respectively for the original data set. The ROC curves for the alternate data set with artificial clutter are shown in Figures 4.45 and 4.46 for the equirectangular and Mercator projections respectively.

In all cases, the 2D versions of the algorithm present lower performances than 3D versions with probabilities of detection ranging between 50% and 70% for small values of FPR. Furthermore, in both versions of the data set, the equirectangular projection exhibits slightly better results than the Mercator projection.

Table 4.8 ranks the methods according to their performance in the ROC curves for values of FPR < 0.1.

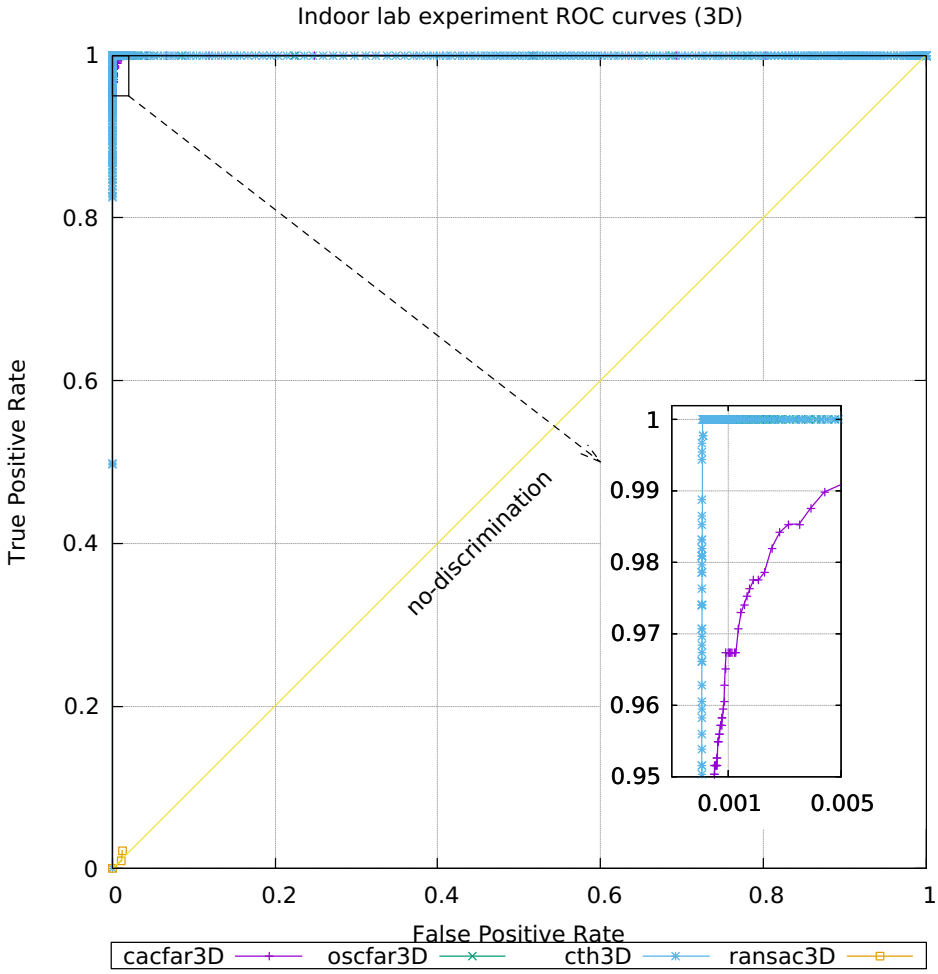

FIGURE 4.41: ROC curves of the methods applied to 3D PCD in the indoor laboratory experimental data set.

In all cases the OS-CFAR adaptive version presented the highest performance. Also, in all situations, the adaptive CFAR version performed better than the standard counterpart. Both, adaptive and standard, OS-CFAR versions showed a better performance than the segmentation thresholding method. As in the results for 3D PCD, the CA-CFAR methods presented a lower performance than the OS-CFAR methods and in most cases than the segmentation thresholding method.

Detection probability and actual false alarm rate versus required probability of false alarm for both, original and alternate with clutter, versions of the data set are presented in Figures 4.47, 4.48, 4.49 and 4.50, respectively.

Table 4.9 show the methods ranked in terms of their detection probability and false alarm rate for each version of the data set.

These results are consistent with the results of the ROC curves, showing that the 3D

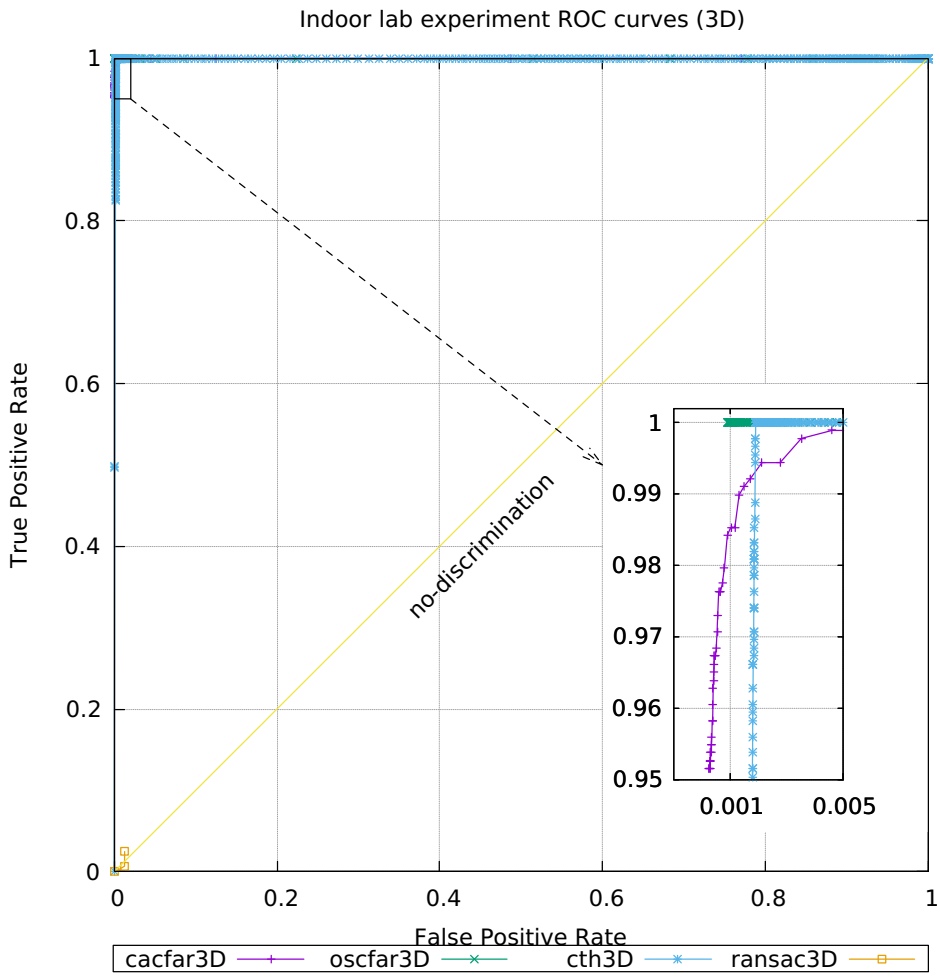

FIGURE 4.42: ROC curves of the methods applied to 3D PCD in the indoor laboratory experimental data set with artificial clutter.

versions of the methods obtain a much higher TPR. In terms of the actual false alarm rate, all methods present a low performance for required false alarm rate values under 1×10^{-3}. However, CA-CFAR 3D presents the best actual false alarm rate in both versions of the data set.

The detection probability achieved by the methods applied to 2D projected data are in the range between 52% and 62% for required values of FPR < 0.1, in the original data set. In the alternate data set, the detection probabilities are lower in the range between 50% an 60% for values of FPR < 0.1 The same behaviour is observed with respect to the actual false alarm rate obtained by the 2D version of the methods.

The OSPA metric results are presented in Figures 4.51, 4.52 and 4.53 for 2D projected data and 3D PCD. Respectively, Figures 4.54, 4.55 and 4.56 for 2D projected data and 3D PCD show OSPA metric results for the high clutter version of the dataset. The conditions

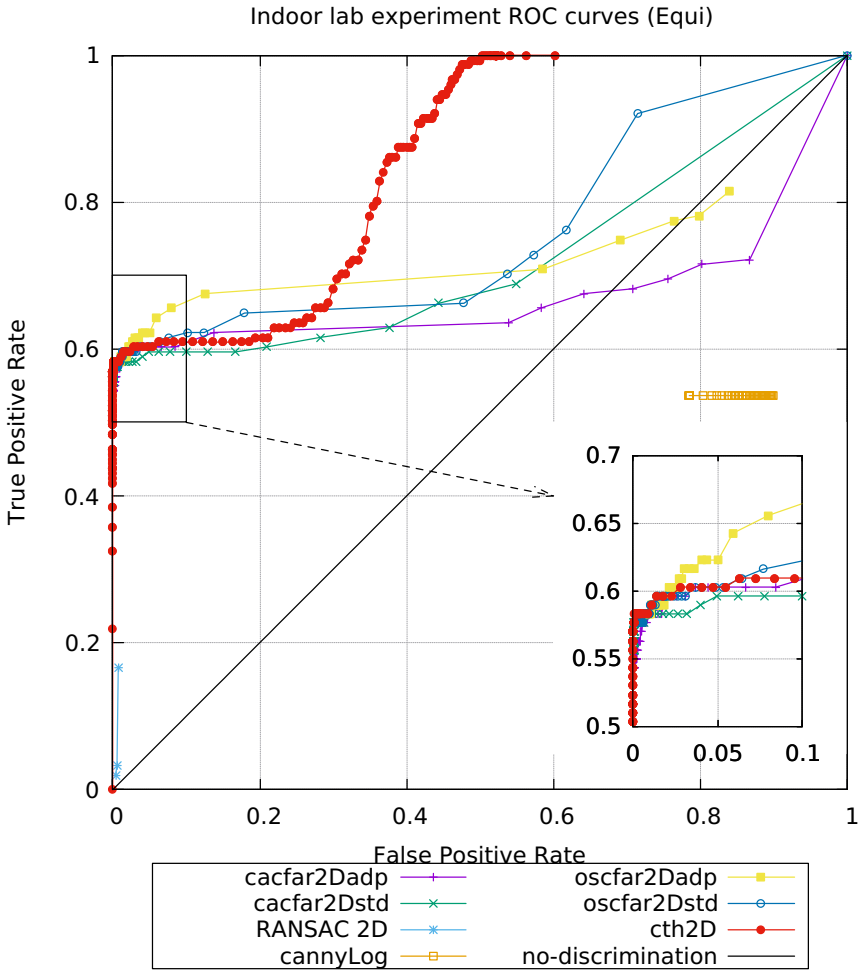

FIGURE 4.43: ROC curves of the methods applied to equirectangular 2D projected data in the indoor laboratory experimental data set.

are the same as the one described in previous results sections of the simulated data sets.

For TPR values higher than 0.6 in 2D projected data (both normal and high clutter versions), all methods present comparable OSPA error curves, dominated by the high FPR values. Segmentation threshold, on the other hand present lower OSPA errors for TPR $<$ 0.6, followed by CA-CFAR adaptive version in the high clutter dataset, and by CA-CFAR standard in the original dataset.

Consistent with the detection statistics results, OS-CFAR 3D presents a lower OSPA error in the high clutter 3D PCD environment. CA-CFAR 3D yielded the lowest OSPA errors for TPR values between 0.95 and 0.99 in both, normal and high clutter versions of the data.

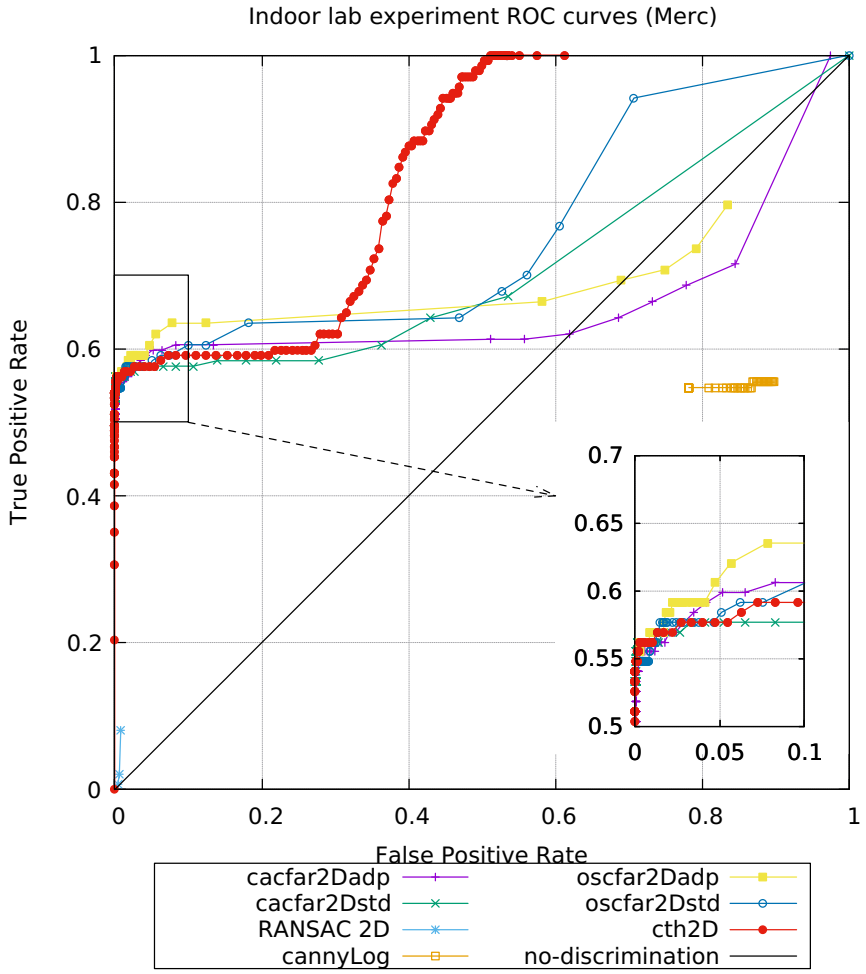

FIGURE 4.44: ROC curves of the methods applied to Mercator 2D projected data in the indoor laboratory experimental data set.

TABLE 4.8: Ranking of 2D methods based on ROC curves for FPR < 0.1, "Original" refers to the original data set without additional clutter, while the "Alternate" corresponds to a copy of the data set with patch of added clutter points.

Rank	Original (Equi)	Original (Merc)	Alternate (Equi)	Alternate (Merc)
1	oscfar2Dadp	oscfar2Dadp	oscfar2Dadp	oscfar2Dadp
2	oscfar2Dstd	cacfar2Dadp	oscfar2Dstd	oscfar2Dstd
3	cth2D	oscfar2Dstd	cth2D	cth2D
4	cacfar2Dadp	cth2D	cacfar2Dadp	cacfar2Dadp
5	cacfar2Dstd	cacfar2Dstd	cacfar2Dstd	cacfar2Dstd
6	RANSAC 2D	RANSAC 2D	RANSAC 2D	RANSAC 2D
7	cannyLog	cannyLog	cannyLog	cannyLog

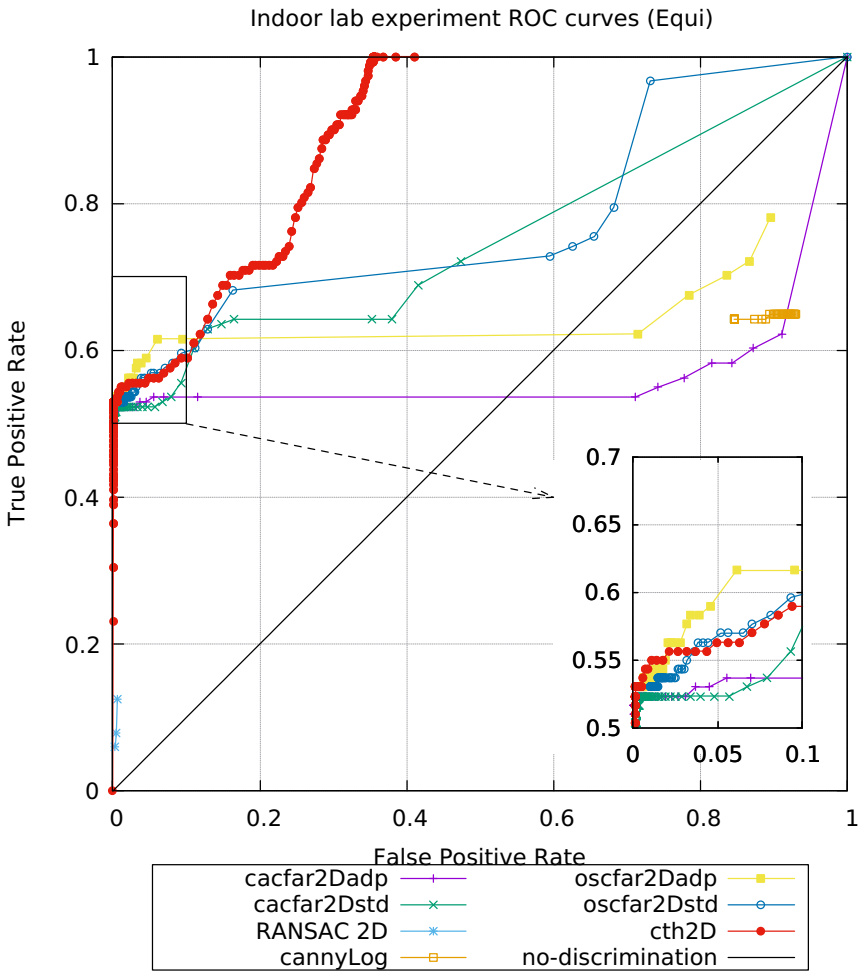

FIGURE 4.45: ROC curves of the methods applied to equirectangular 2D projected data in the indoor laboratory experimental data set (artificial clutter).

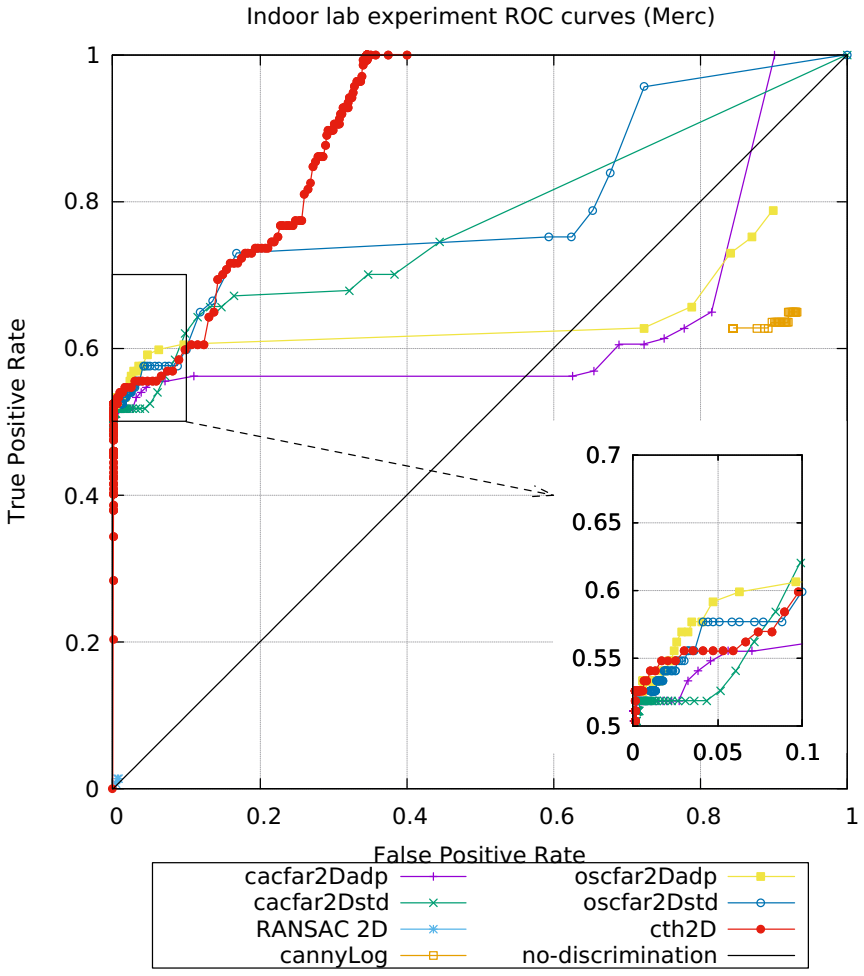

FIGURE 4.46: ROC curves of the methods applied to Mercator 2D projected data in the indoor laboratory experimental data set (artificial clutter).

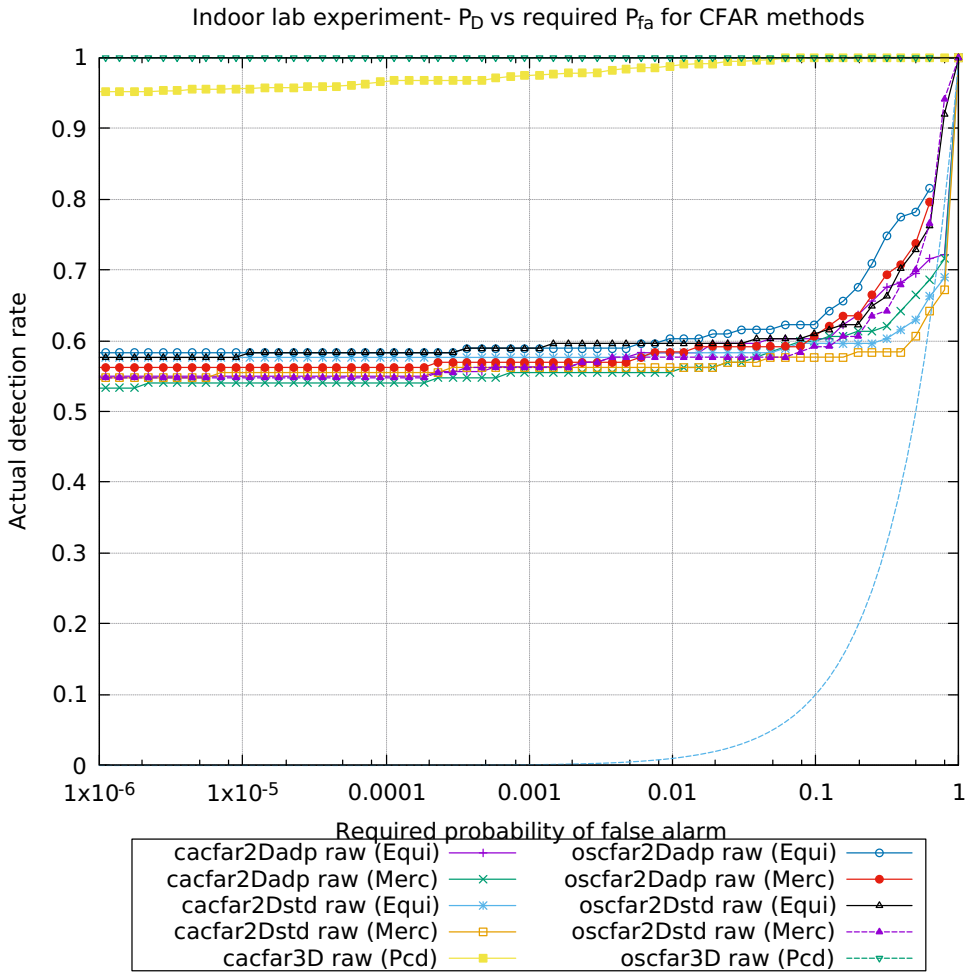

FIGURE 4.47: Probability of detection vs. required false alarm rate for indoor laboratory data.

TABLE 4.9: Ranking of CFAR methods based on probability of detection (TPR) and actual false alarm rate (FPR).

Rank	Original TPR	Original actual FPR	Alternate TPR	Alternate actual FPR
1	oscfar3D PCD	cacfar3D PCD	oscfar3D PCD	cacfar3D PCD
2	cacfar3D PCD	cacfar2Dstd (Merc)	cacfar3D PCD	cacfar2Dadp (Merc)
3	oscfar2Dadp (Equi)	cacfar2Dstd (Equi)	oscfar2Dstd (Merc)	cacfar2Dadp (Equi)
4	oscfar2Dstd (Equi)	cacfar2Dadp (Merc)	oscfar2Dadp (Equi)	oscfar3D PCD
5	cacfar2Dstd (Equi)	cacfar2Dadp (Equi)	oscfar2Dstd (Equi)	cacfar2Dstd (Equi)
6	oscfar2Dadp (Merc)	oscfar3D PCD	oscfar2Dadp (Merc)	cacfar2Dstd (Merc)
7	cacfar2Dstd (Merc)	oscfar2Dstd (Merc)	cacfar2Dstd (Equi)	oscfar2Dadp (Merc)
8	cacfar2Dadp (Equi)	oscfar2Dadp (Merc)	cacfar2Dstd (Merc)	oscfar2Dadp (Equi)
9	oscfar2Dstd (Merc)	oscfar2Dstd (Equi)	cacfar2Dadp (Equi)	oscfar2Dstd (Equi)
10	cacfar2Dadp (Merc)	oscfar2Dadp (Equi)	cacfar2Dadp (Merc)	oscfar2Dstd (Merc)

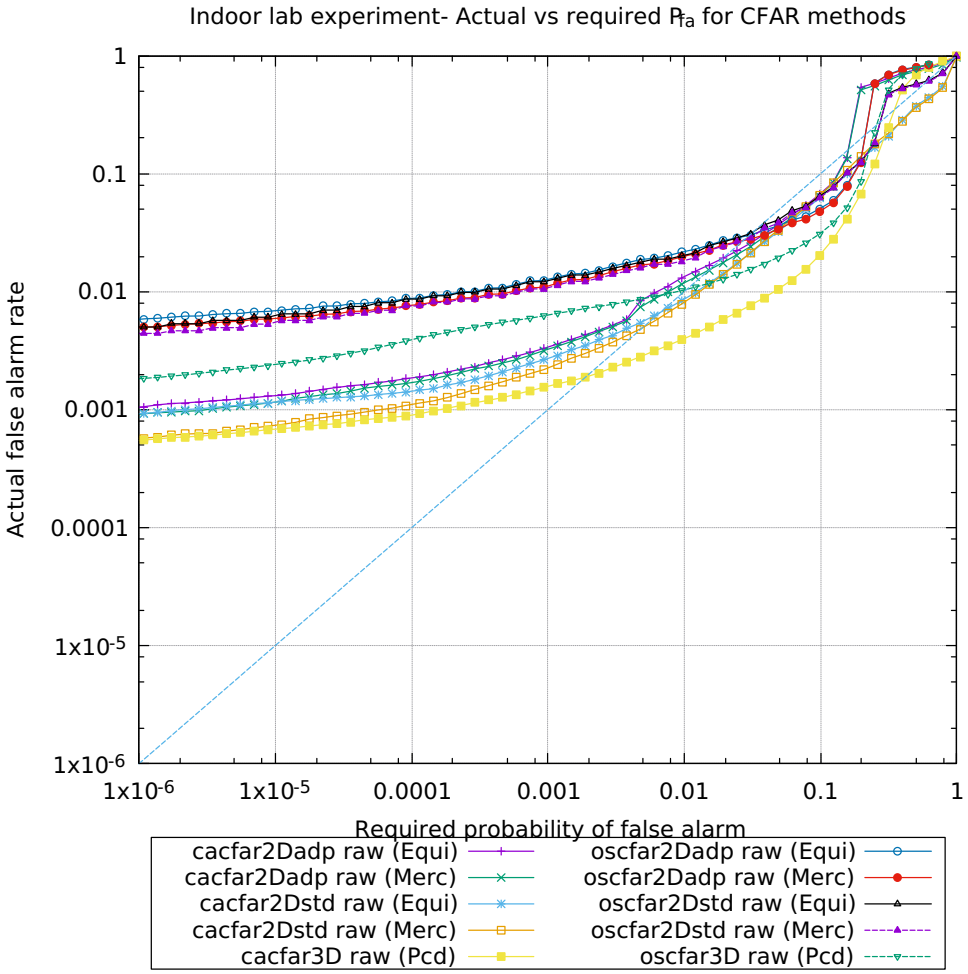

FIGURE 4.48: Probability of false alarm vs. required false alarm rate for indoor laboratory data.

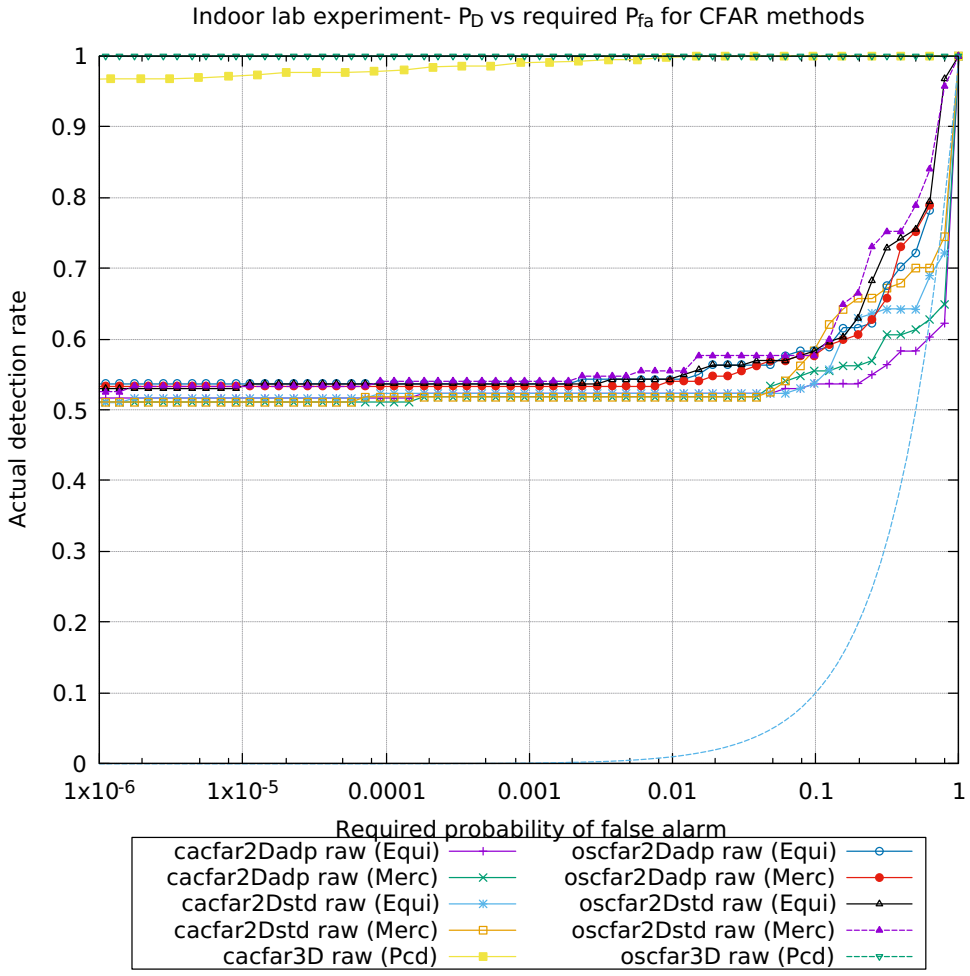

FIGURE 4.49: Probability of detection vs. required false alarm rate for indoor laboratory data (artificial clutter).

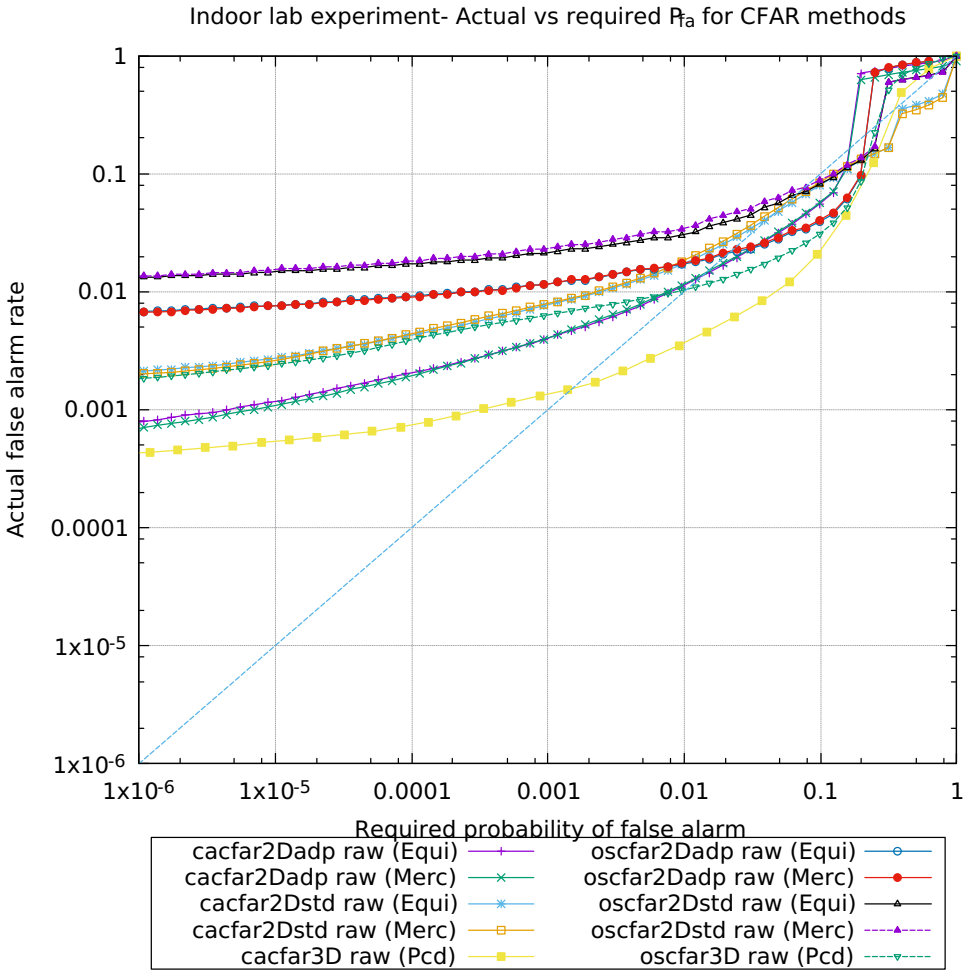

FIGURE 4.50: Probability of false alarm vs. required false alarm rate for indoor laboratory data (artificial clutter).

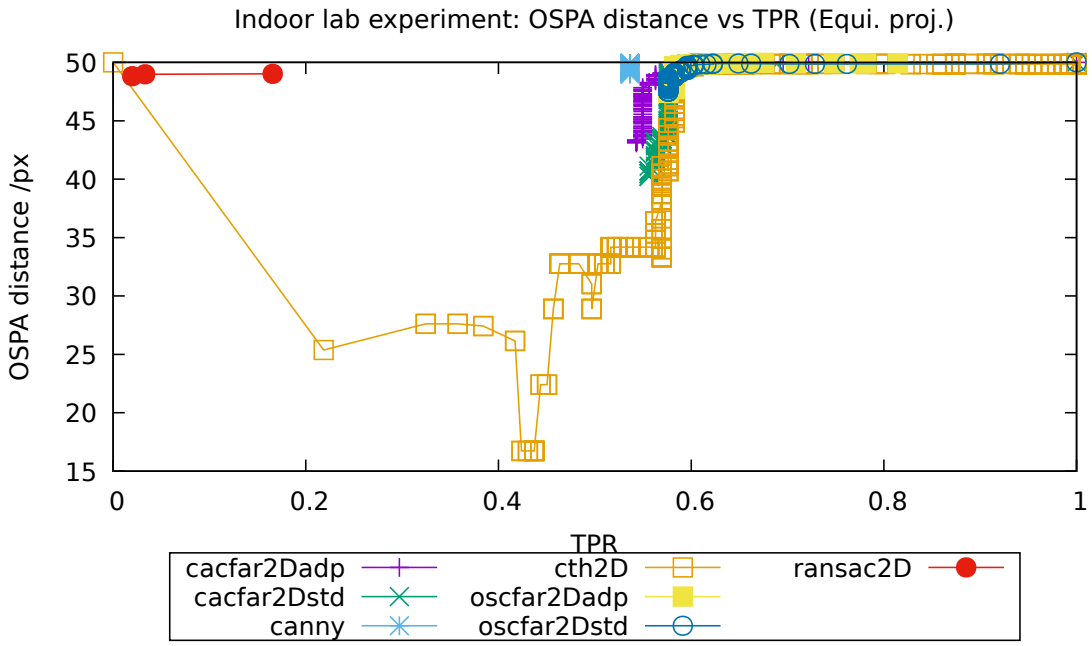

FIGURE 4.51: OSPA metric results for indoor lab experiment (Equirectangular projection).

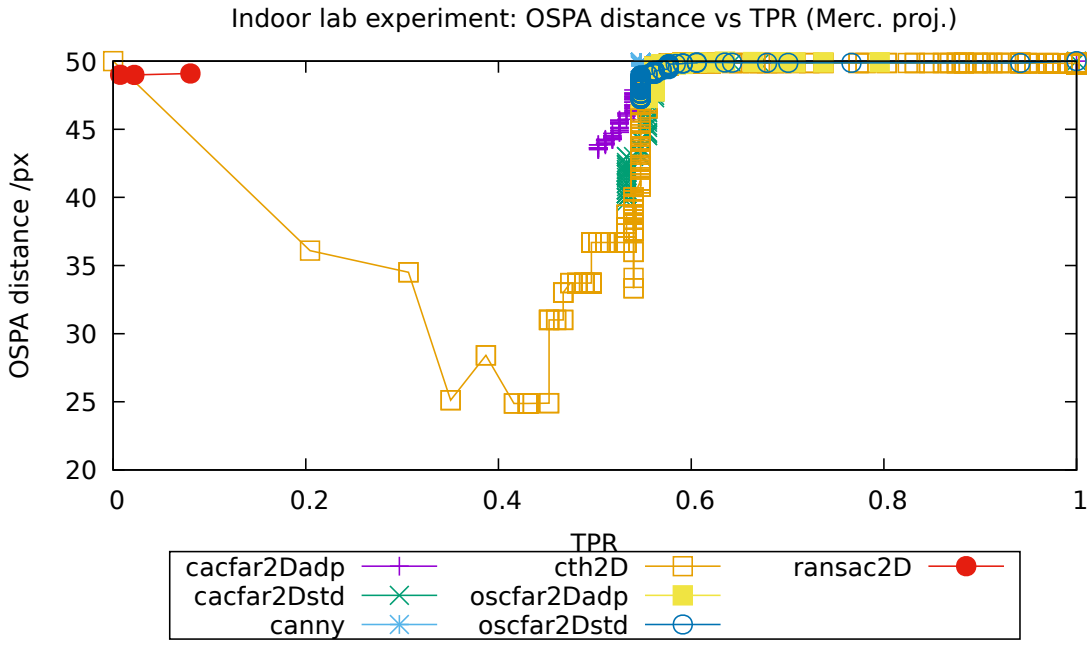

FIGURE 4.52: OSPA metric results for indoor lab experiment (Mercator projection).

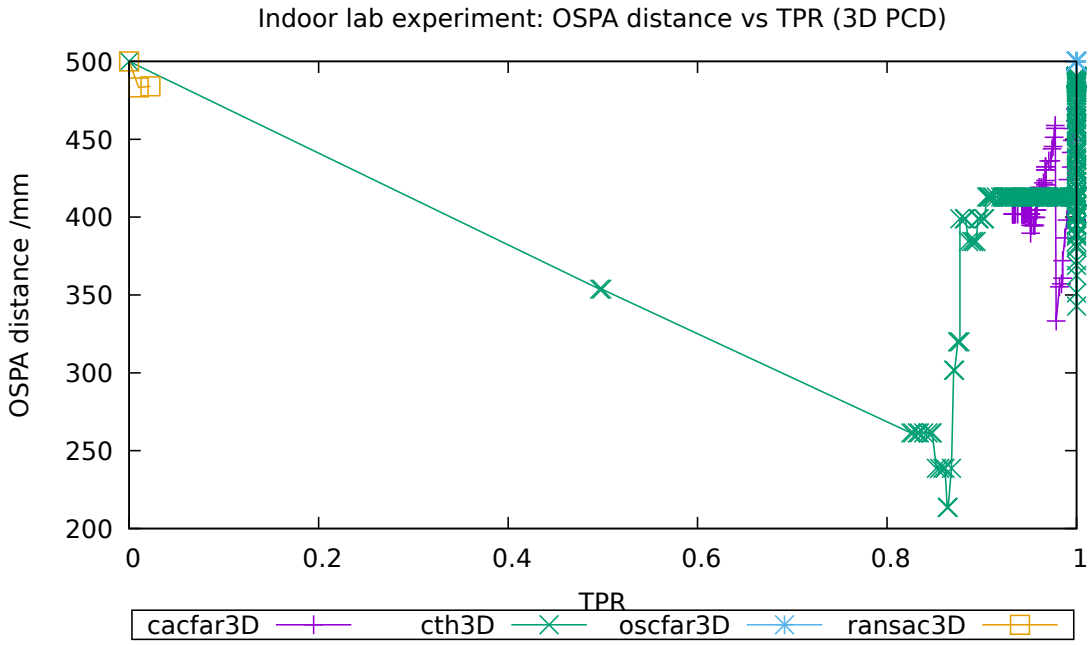

FIGURE 4.53: OSPA metric results for indoor lab experiment (3D PCD).

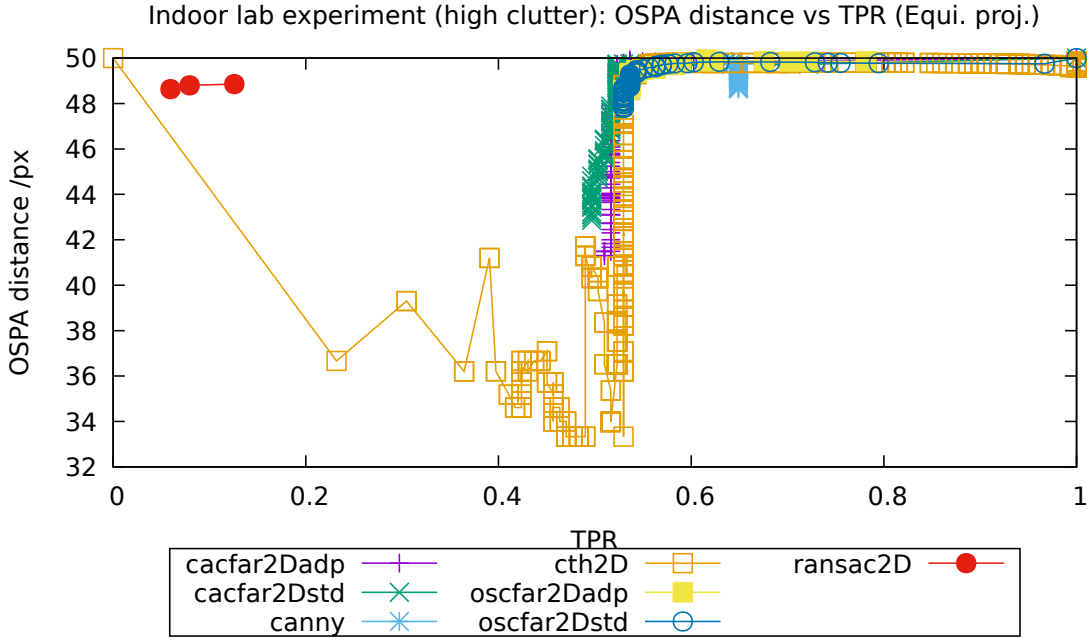

FIGURE 4.54: OSPA metric results for indoor lab experiment with high clutter (Equirectangular projection).

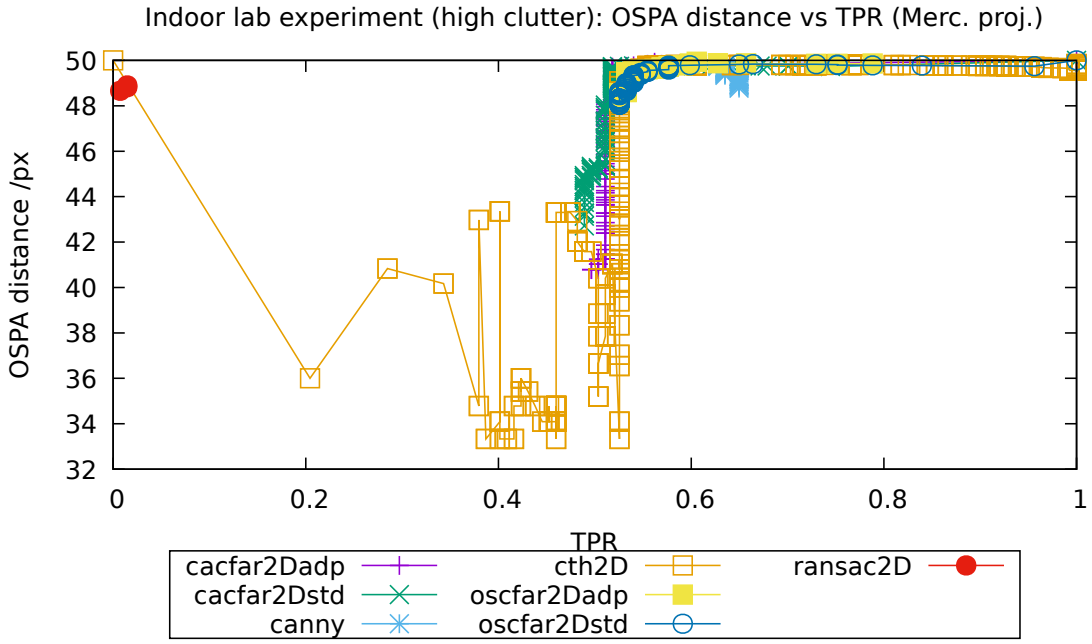

FIGURE 4.55: OSPA metric results for indoor lab experiment with high clutter (Mercator projection).

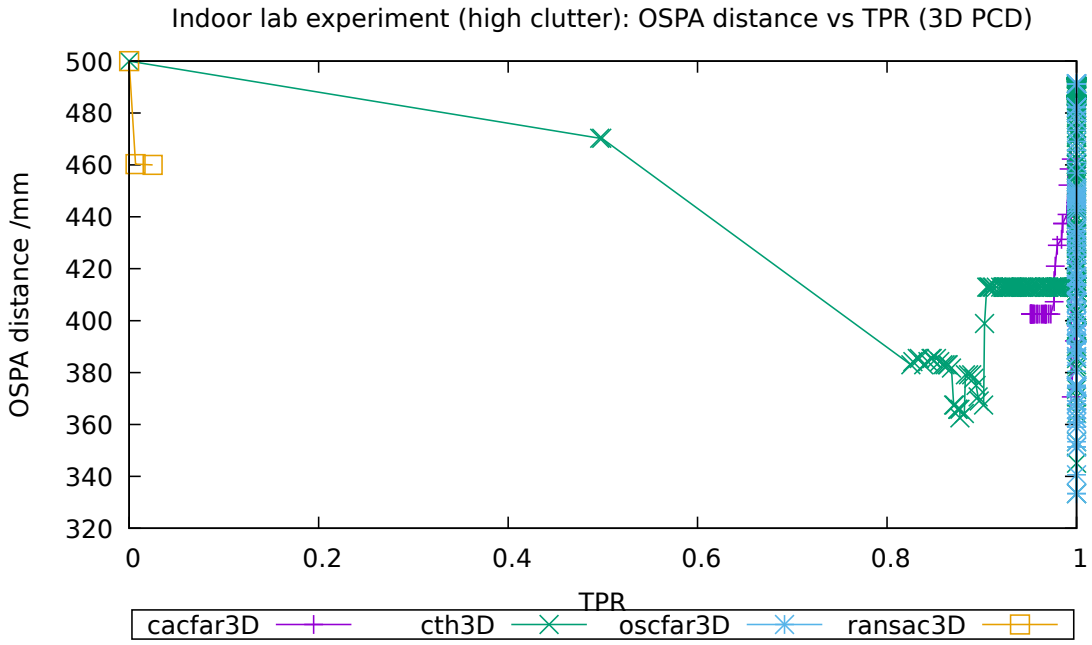

FIGURE 4.56: OSPA metric results for indoor lab experiment with high clutter (3D PCD).

4.8 Mine environment test

A data-set obtained in an underground mine has also been used to test the algorithms. The data-set contains a large number of points, therefore it was partitioned into two sections of the mine's tunnel. The first one (see Figures 4.57, 4.58 and 4.59) near the end of the tunnel with 7 markers, 5 in the tunnel (coloured red in the figure) and 2 (coloured green in the Figure) on the end wall.

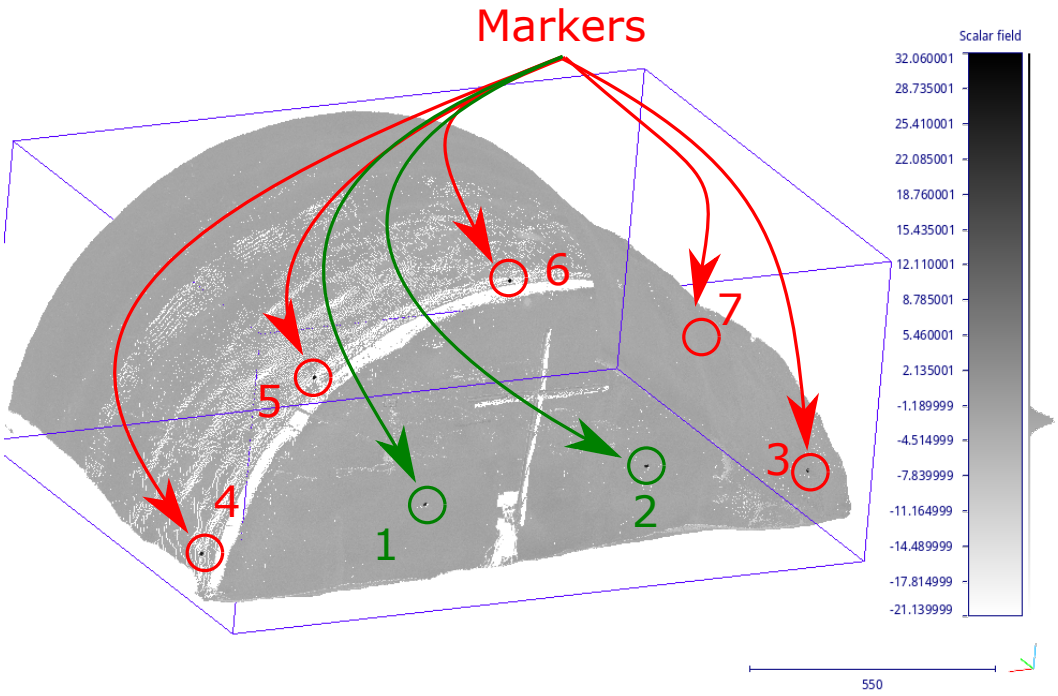

FIGURE 4.57: End section of an underground mine tunnel. Red and green circles show the visible targets' locations.

This section has homogeneous (reflected intensity with a low variance) tunnel background and the targets present a high SNR.

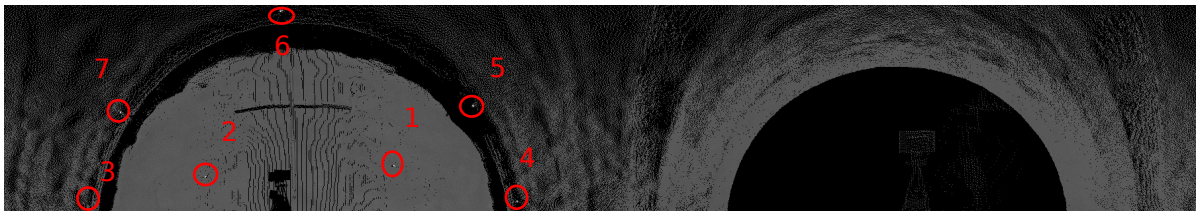

FIGURE 4.58: Mercator projection of end section of mine tunnel.

The second partition correspond to the mid section of the tunnel, where two sets of 5 markers are located across the tunnel at two positions each on one end of the section (Figures 4.60,

4.61 and 4.62, with blue markers are on one end, magenta markers on the other end).

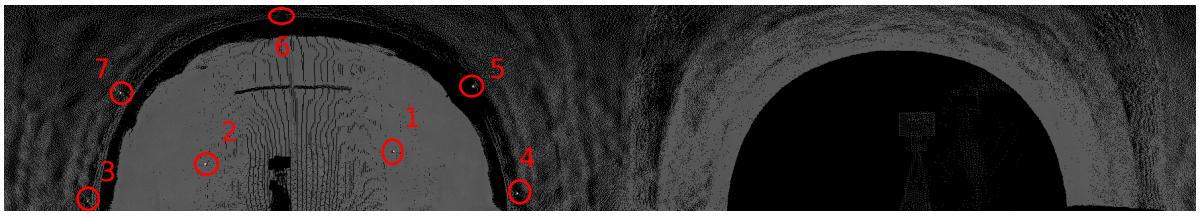

FIGURE 4.59: Equirectangular projection of end section of mine tunnel.

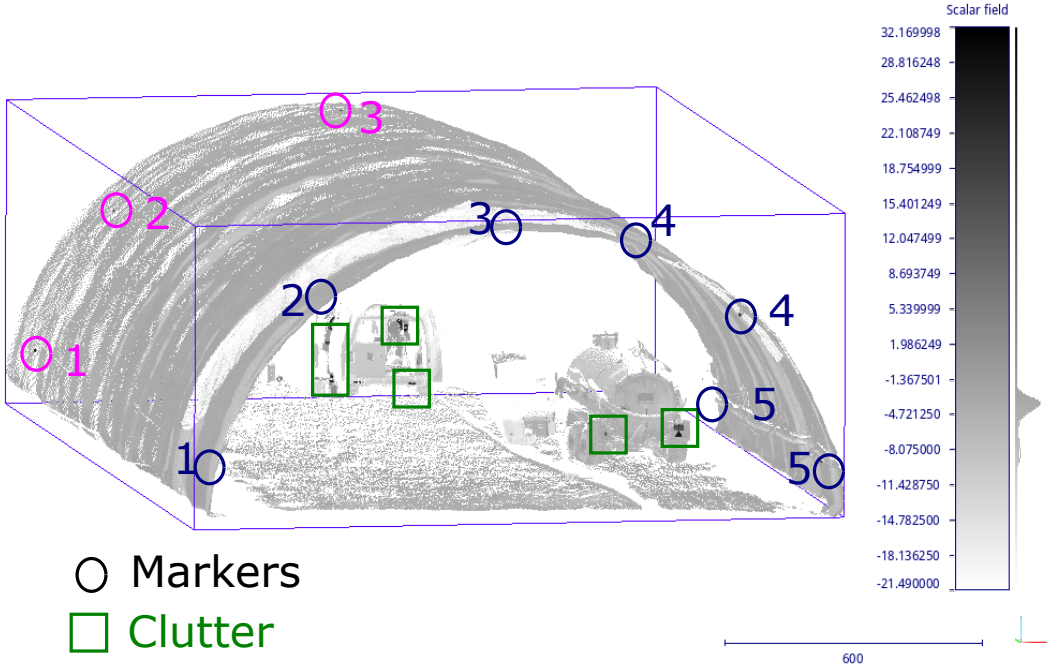

FIGURE 4.60: Mid section of an underground mine tunnel. Blue and magenta circles show the visible targets' locations, green squares enclose visible clutter objects.

As in the indoor laboratory data set, for each section of the mine, an alternate data set consisting of a copy of the original data with a small patch incorporating artificial clutter is used to analyse the effect of higher clutter level on the different methods.

FIGURE 4.61: Mercator projection of mid section of mine tunnel.

Also, the tunnel background is non-homogeneous in some sections increasing the noise.

FIGURE 4.62: Equirectangular projection of mid section of mine tunnel.

4.8.1 First section results

Figures 4.63 and 4.64 present the ROC curves of the methods applied to 3D PCD for the original an the alternate data sets, respectively.

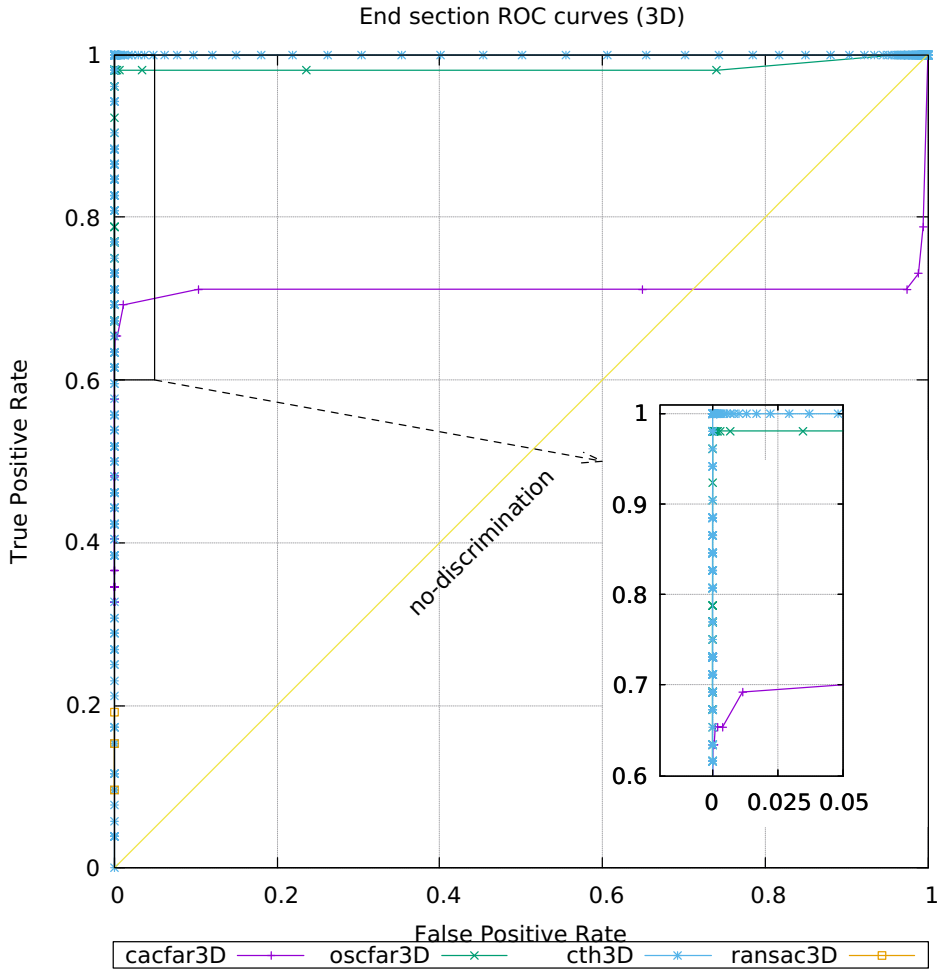

FIGURE 4.63: ROC curves of the methods applied to 3D PCD in the end section of underground mine tunnel data set.

The same behaviour described in the previous section for the indoor laboratory data set is observed in the plots. The CFAR 3D methods are not affected by the higher level of clutter in the alternate data set, while the segmentation thresholding method increases its

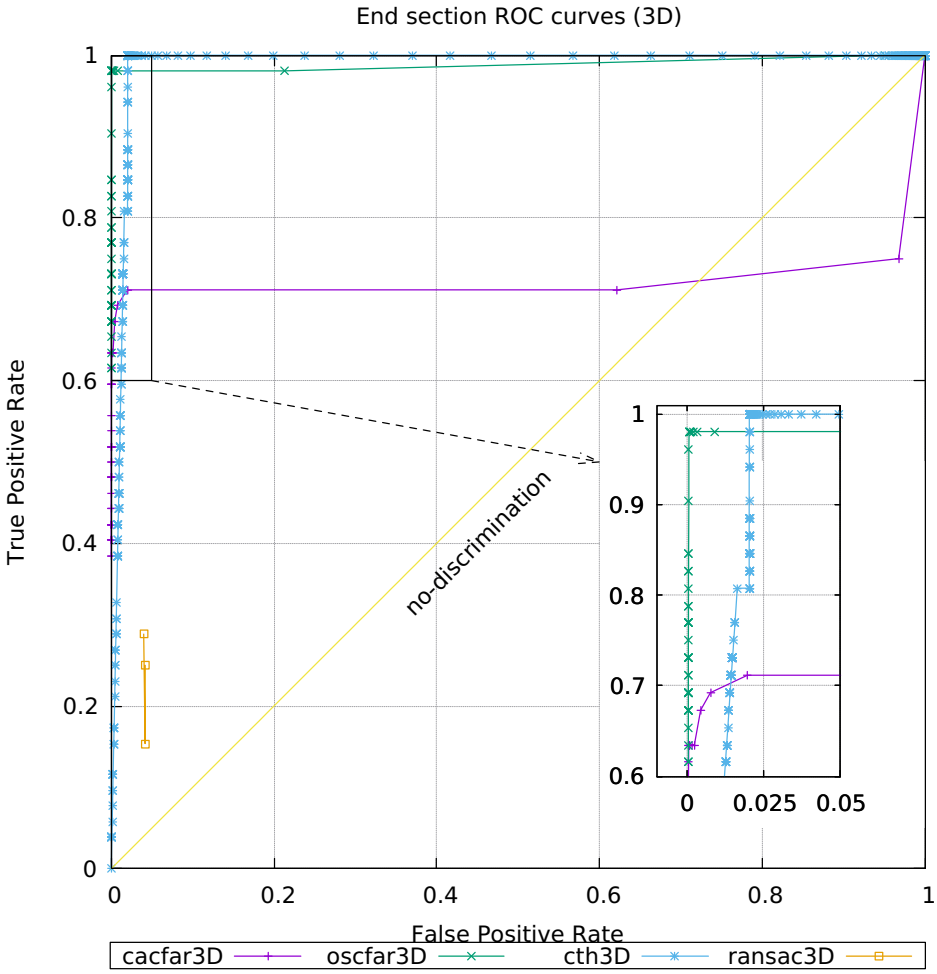

FIGURE 4.64: ROC curves of the methods applied to 3D PCD in the end section of underground mine tunnel data set with artificial clutter.

false alarm rate. The CA-CFAR 3D method improves slightly, however in this data set, CA-CFAR presents a much lower probability of detection than the other methods (near 70% for most values of FPR). OS-CFAR 3D presents a very high detection rate, nevertheless it is slightly below the segmentation threshold method in terms of the ROC curve.

ROC curves for 2D projected data of the original data set are presented in Figures 4.65 and 4.66 (equirectangular and Mercator projections, respectively), while the ROC curves of the alternate data set are shown in Figures 4.67 and 4.68 (equirectangular and Mercator projections, respectively).

As in the previous data sets, equirectangular projection yields, in general, relatively higher performances than the Mercator projection. However, unlike previous results, the alternate data set with an artificial clutter patch presents slightly better performances than the original one. Nevertheless, most methods exhibit similar levels of performance.

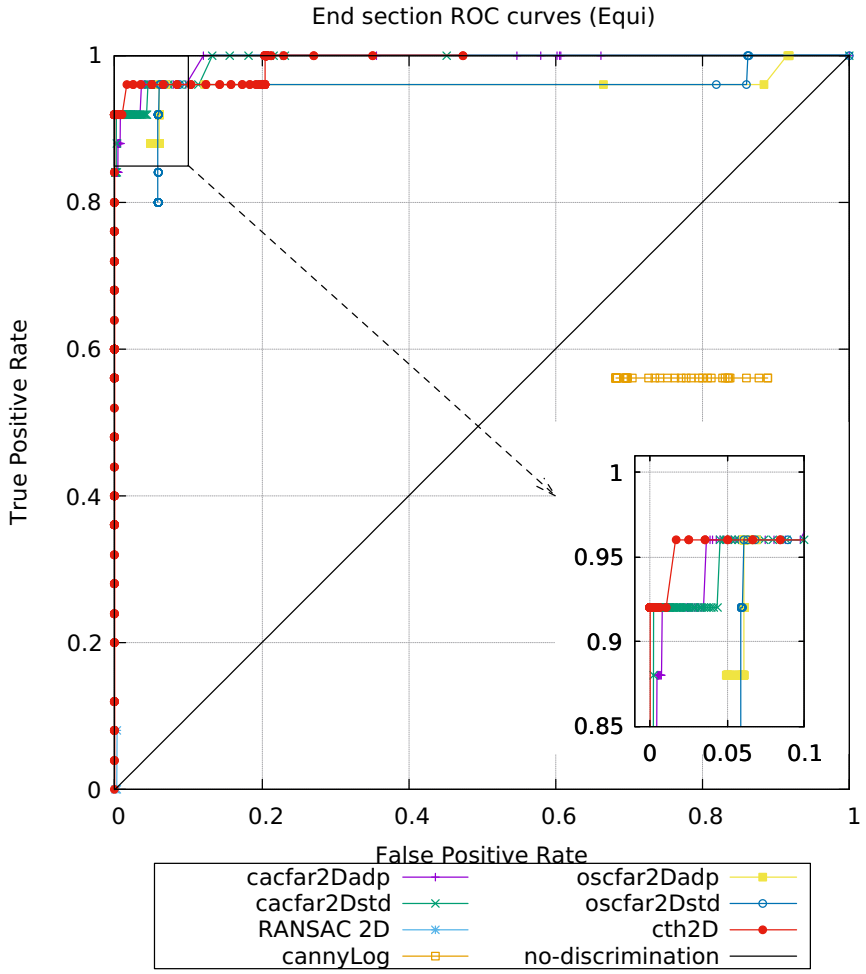

FIGURE 4.65: ROC curves of the methods applied to equirectangular 2D projected data in the end section of underground mine tunnel data set.

The 2D versions of the methods are ranked in table 4.10 based on the ROC curves results.

In original data set, segmentation threshold is ranked first. The segmentation threshold method shows a slightly higher FPR in the alternate data set, which allows a CA-CFAR method to rank first, for very small values of FPR. In all cases, CA-CFAR methods presented better performances than the OS-CFAR methods. In all but the equirectangular projection of the alternate data set, the CFAR adaptive versions where ranked higher than their standard counterpart.

Figures 4.69 and 4.70 show probability of detection and actual probability of false alarm, respectively, for the original data set. The same plots for the alternate data set are presented in Figures 4.71 and 4.72.

The alternate data set presents moderately higher performances than the original data set in terms of detection probability, while all the methods, except CA-CFAR 2D standard,

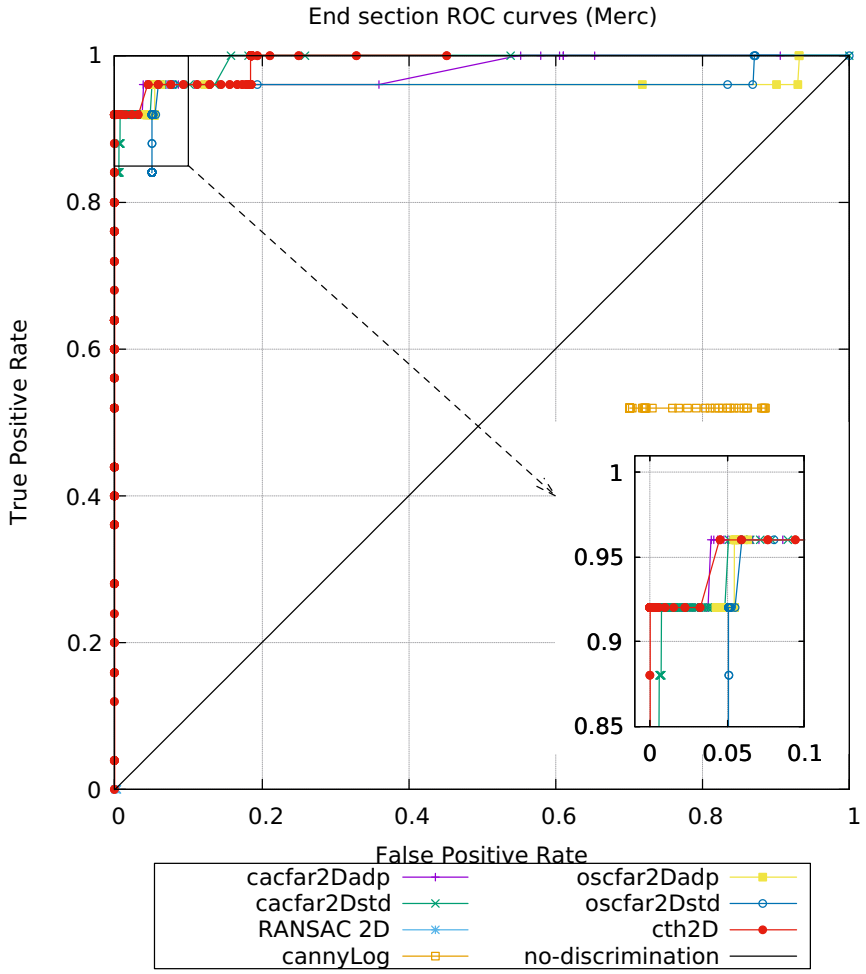

FIGURE 4.66: ROC curves of the methods applied to Mercator 2D projected data in the end section of underground mine tunnel data set.

show comparable performances with respect to actual false alarm rate in both versions of the data set.

The methods are ranked in table 4.11 based on the detection probability and actual false alarm rate results for each version of the data set.

Overall, OS-CFAR 3D is the only method showing a high detection probability (greater than 70% in the whole range of required false alarm probability), which simultaneously presents an actual false alarm rate similar or below the required value.

The 2D adaptive versions of the CFAR methods presented a good performance in terms of detection probability. CA-CFAR 2D version also show, better probabilities of detection than OS-CFAR 2D, in general. The 2D adaptive CA-CFAR methods are ranked at the top, among the 2D CFAR versions, with respect to the actual false alarm rate.

The OSPA metric results are shown in Figures 4.73, 4.74 and 4.75 (original clutter ver-

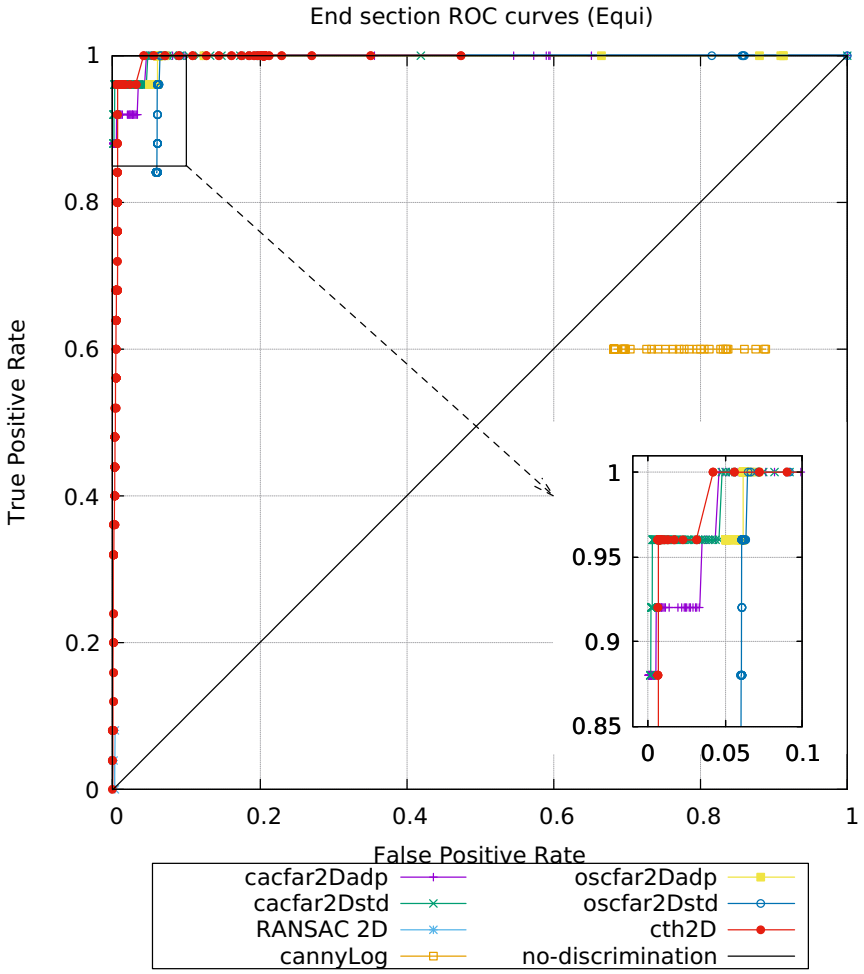

FIGURE 4.67: ROC curves of the methods applied to equirectangular 2D projected data in the end section of underground mine tunnel data set (artificial clutter).

sion), as well as, in Figures 4.76, 4.77 and 4.78 (high clutter version).

As expected, due to the high SNR, the segmentation threshold method achieves lower OSPA errors. However, in the artificial clutter version of the data, all methods present high OSPA errors. Also, CA-CFAR adaptive version in equirectangular 2D projected data shows an OSPA error comparable to the segmentation threshold for TPR values around 0.9, while CA-CFAR 3D and OS-CFAR 3D present lower OSPA errors for TPR values in the range between 0.5 and 0.7, consistent with the ROC curves presented above.

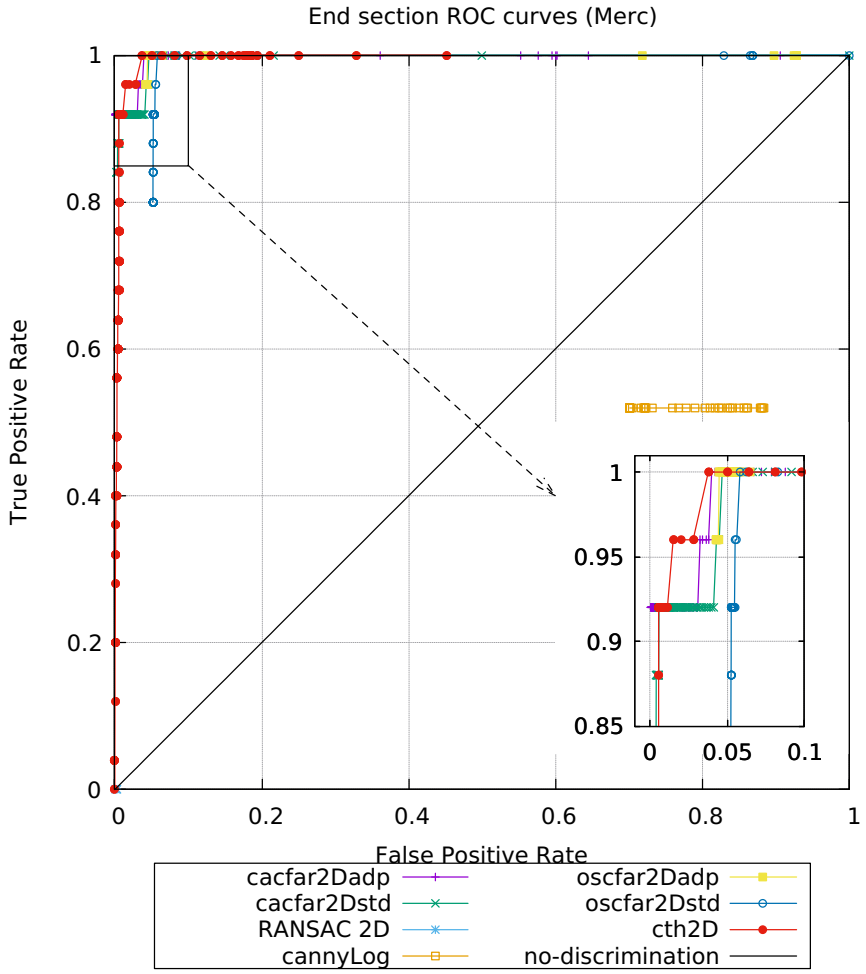

FIGURE 4.68: ROC curves of the methods applied to Mercator 2D projected data in the end section of underground mine tunnel data set (artificial clutter).

TABLE 4.10: Ranking of 2D methods based on ROC curves for FPR < 0.1, "Original" refers to the original data set without additional clutter, while the "Alternate" corresponds to a copy of the data set with patch of added clutter points.

Rank	Original (Equi)	Original (Merc)	Alternate (Equi)	Alternate (Merc)
1	cth2D	cth2D	cacfar2Dstd	cacfar2Dadp
2	cacfar2Dadp	cacfar2Dadp	cth2D	cth2D
3	cacfar2Dstd	cacfar2Dstd	cacfar2Dadp	cacfar2Dstd
4	oscfar2Dadp	oscfar2Dadp	oscfar2Dadp	oscfar2Dadp
5	oscfar2Dstd	oscfar2Dstd	oscfar2Dstd	oscfar2Dstd
6	RANSAC 2D	RANSAC 2D	RANSAC 2D	RANSAC 2D
7	cannyLog	cannyLog	cannyLog	cannyLog

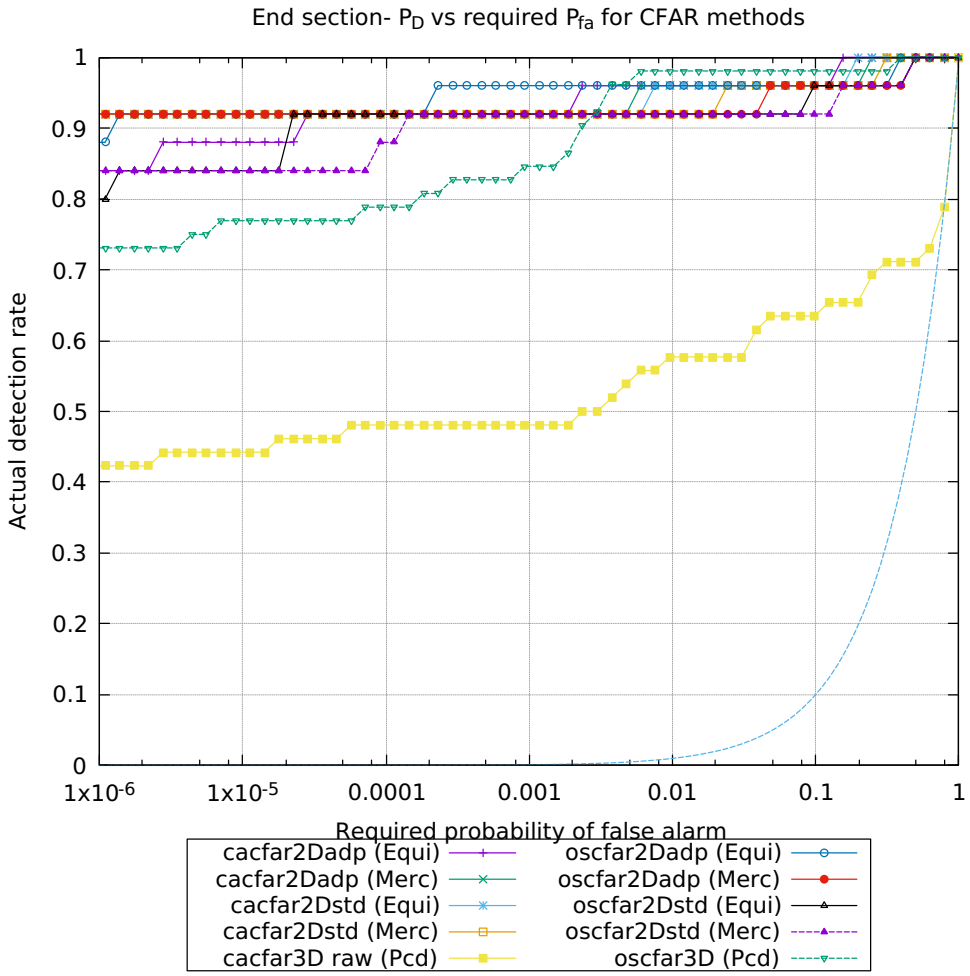

FIGURE 4.69: Probability of detection vs. required false alarm rate for end section of underground mine tunnel data.

TABLE 4.11: Ranking of CFAR methods based on probability of detection (TPR) and actual false alarm rate (FPR).

Rank	Original TPR	Original actual FPR	Alternate TPR	Alternate actual FPR
1	oscfar2Dadp (Equi)	cacfar3D (Pcd)	oscfar2Dadp (Merc)	cacfar3D (Pcd)
2	cacfar2Dadp (Equi)	oscfar3D (Pcd)	oscfar2Dadp (Equi)	oscfar3D (Pcd)
3	cacfar2Dadp (Merc)	cacfar2Dadp(Merc)	cacfar2Dadp (Equi)	cacfar2Dadp (Merc)
4	cacfar2Dstd (Equi)	cacfar2Dadp (Equi)	cacfar2Dadp (Merc)	cacfar2Dadp (Equi)
5	cacfar2Dstd (Merc)	cacfar2Dstd (Merc)	cacfar2Dstd (Equi)	oscfar2Dstd (Merc)
6	oscfar2Dadp (Merc)	cacfar2Dstd (Equi)	cacfar2Dstd (Merc)	oscfar2Dadp (Merc)
7	oscfar2Dstd (Equi)	oscfar2Dstd (Merc)	oscfar2Dstd (Equi)	oscfar2Dstd (Equi)
8	oscfar2Dstd (Merc)	oscfar2Dadp (Merc)	oscfar2Dstd (Merc)	oscfar2Dadp (Equi)
9	oscfar3D (Pcd)	oscfar2Dstd (Equi)	oscfar3D (Pcd)	cacfar2Dstd (Merc)
10	cacfar3D (Pcd)	oscfar2Dadp (Equi)	cacfar3D (Pcd)	cacfar2Dstd (Equi)

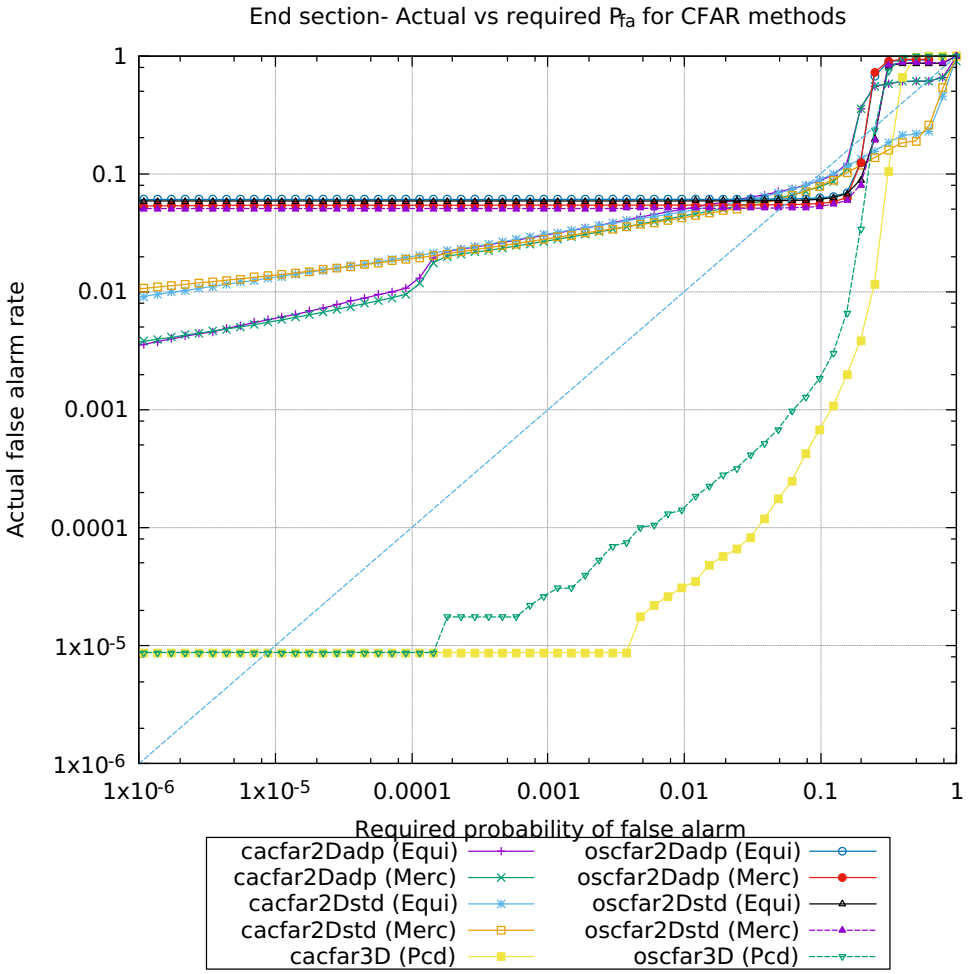

FIGURE 4.70: Probability of false alarm vs. required false alarm rate for end section of underground mine tunnel data.

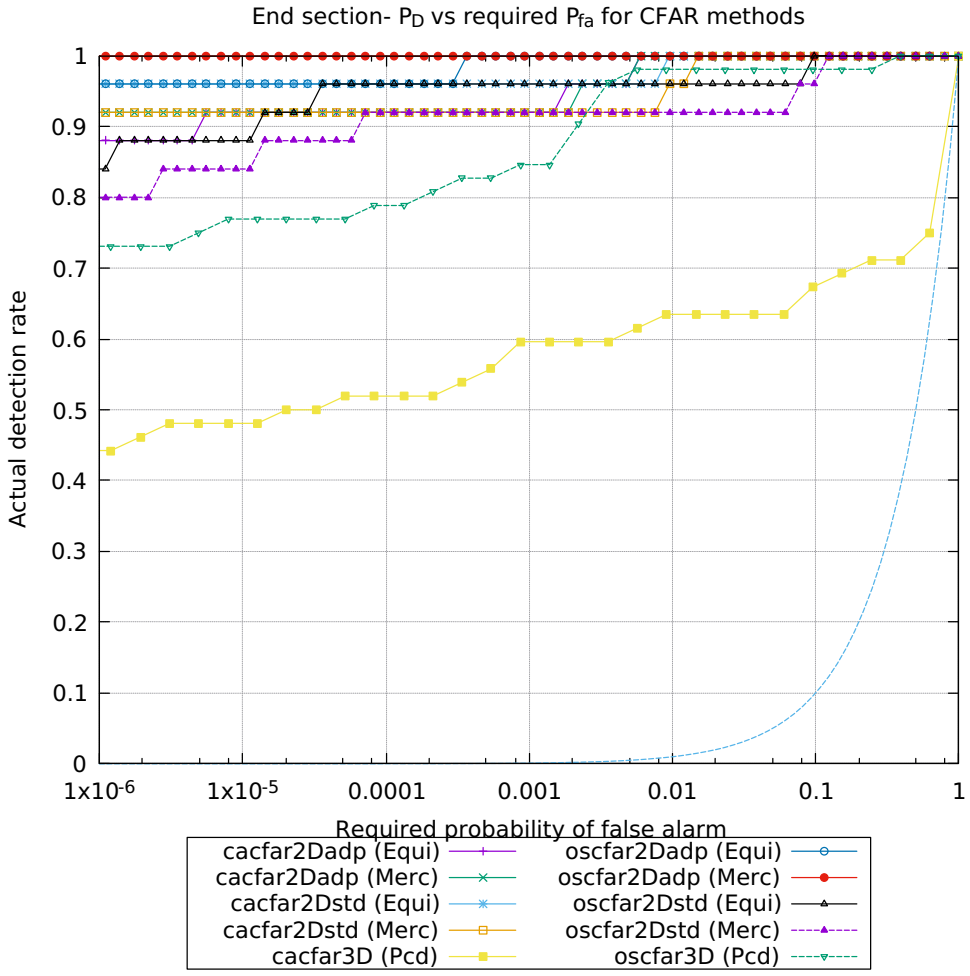

FIGURE 4.71: Probability of detection vs. required false alarm rate for end section of underground mine tunnel data (artificial clutter).

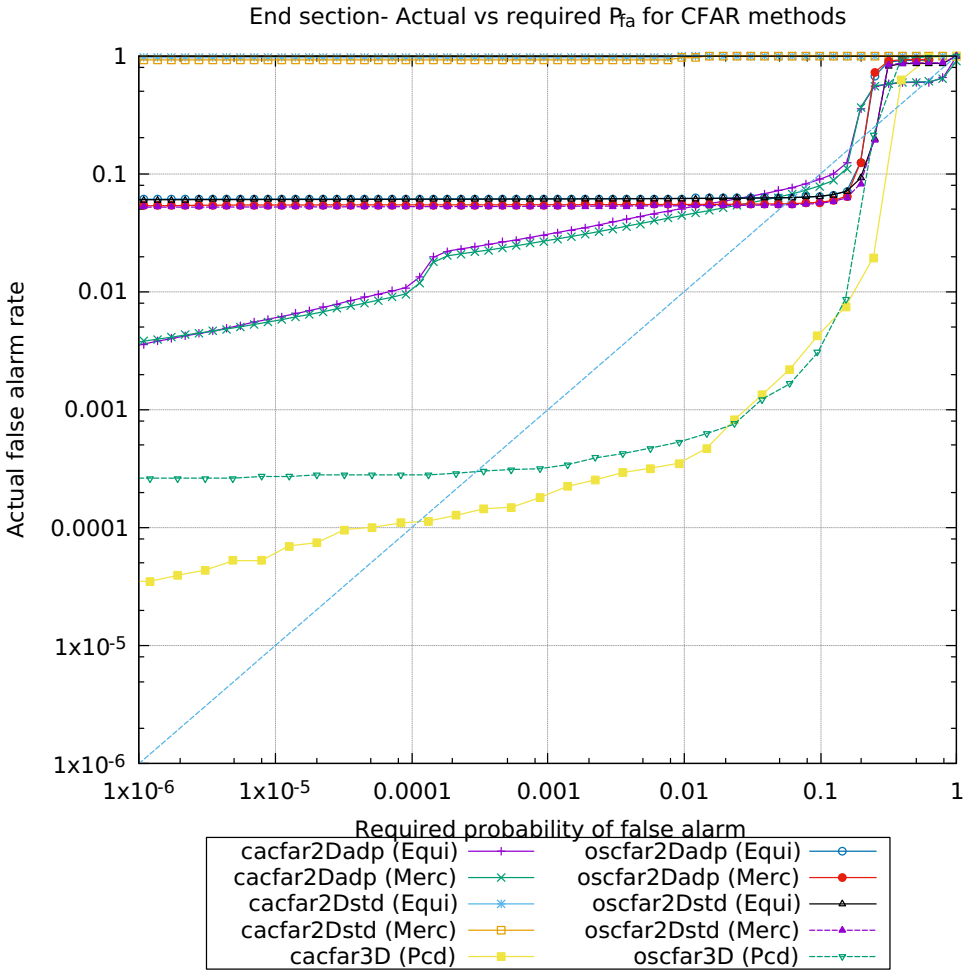

FIGURE 4.72: Probability of false alarm vs. required false alarm rate for end section of underground mine tunnel data (artificial clutter).

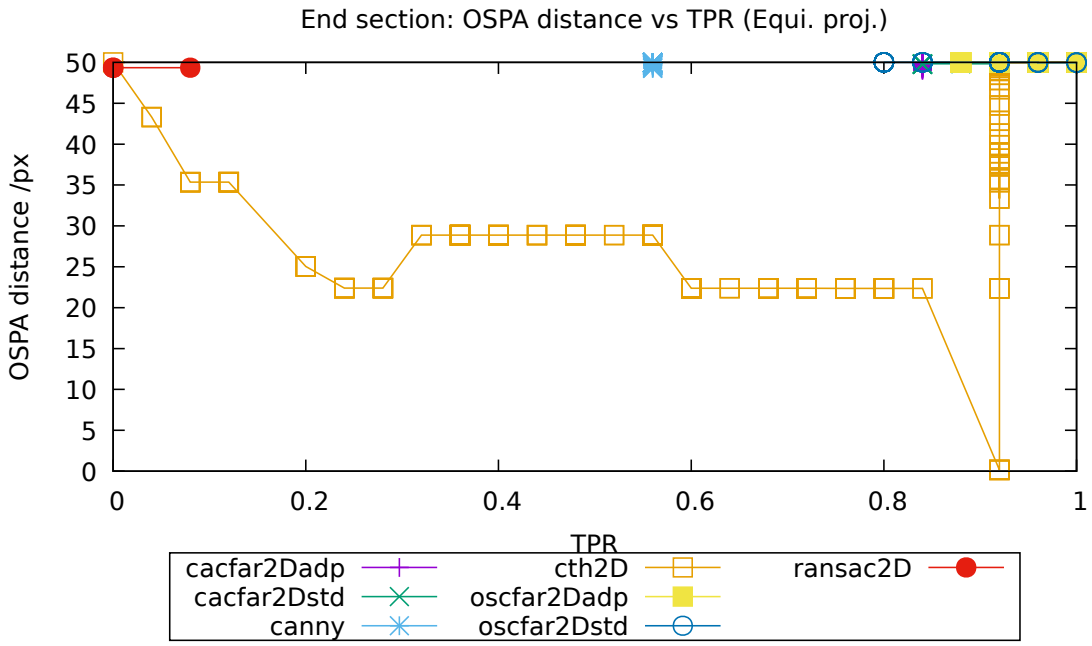

FIGURE 4.73: OSPA metric results for the end section of undergroun mine tunnel (Equirectangular projection).

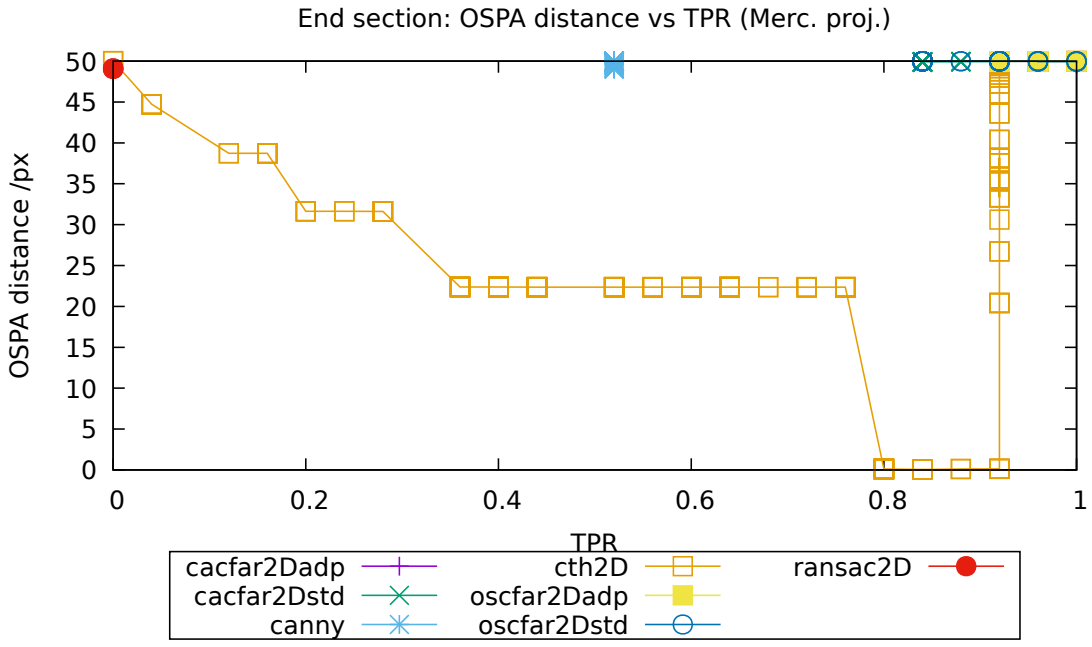

FIGURE 4.74: OSPA metric results for the end section of undergroun mine tunnel (Mercator projection).

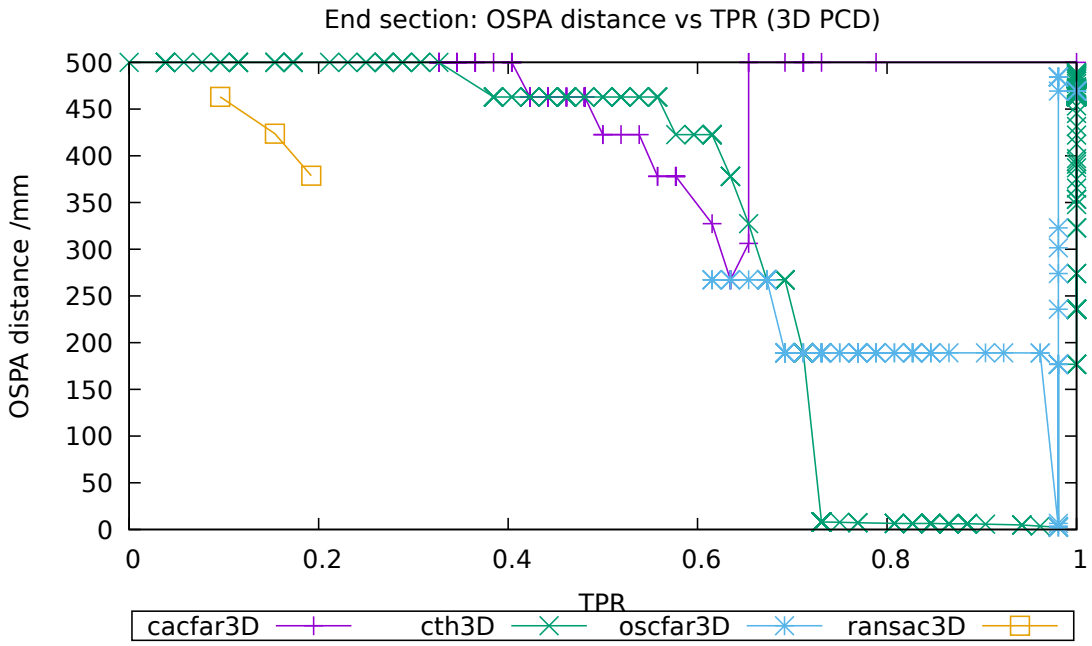

FIGURE 4.75: OSPA metric results for the end section of undergroun mine tunnel (3D PCD).

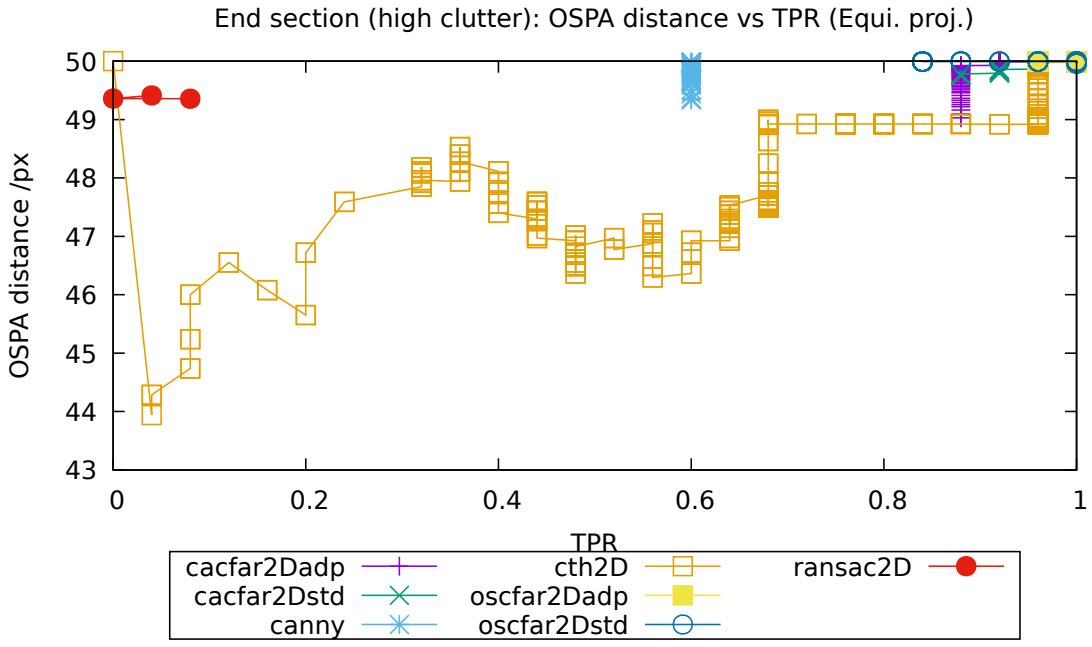

FIGURE 4.76: OSPA metric results for the end section of undergroun mine tunnel with high clutter (Equirectangular projection).

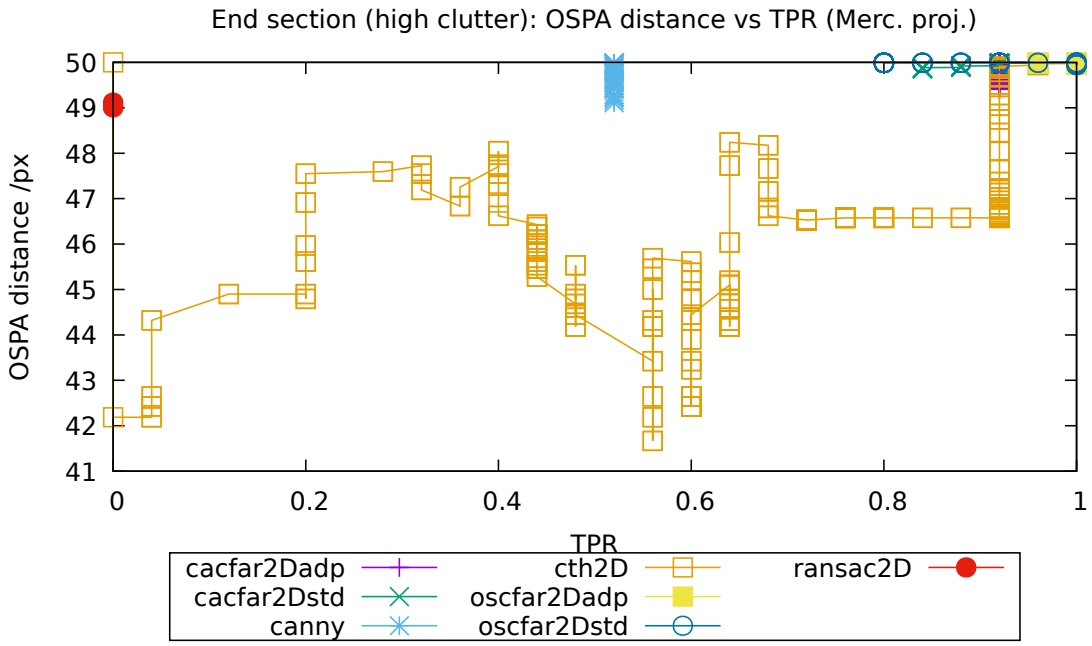

FIGURE 4.77: OSPA metric results for the end section of undergroun mine tunnel with high clutter (Mercator projection).

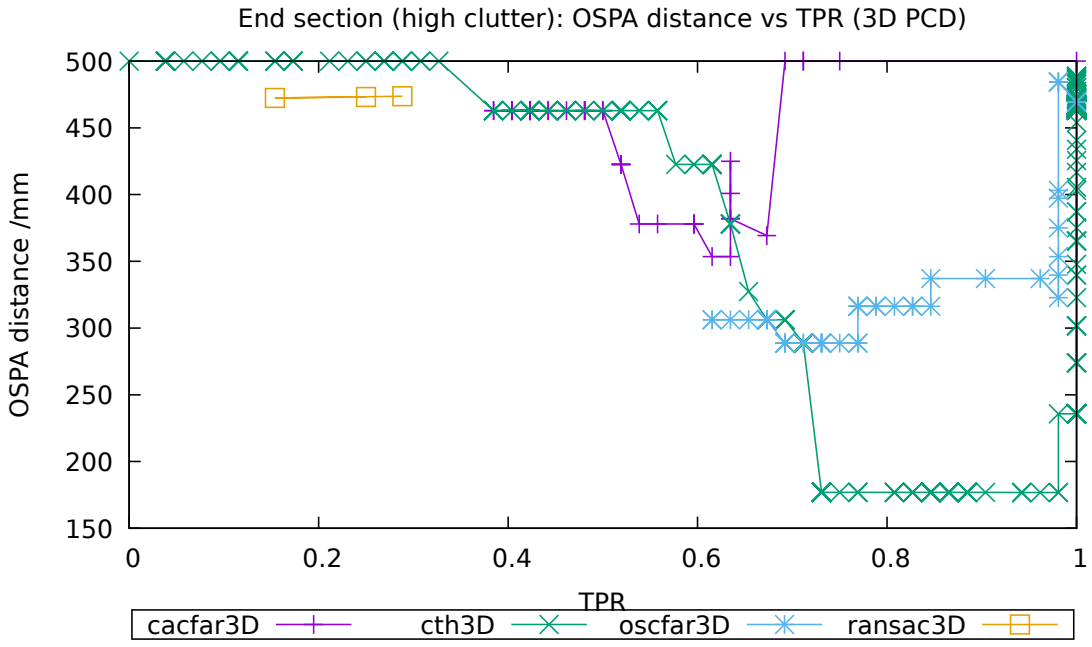

FIGURE 4.78: OSPA metric results for the end section of undergroun mine tunnel with high clutter (3D PCD).

4.8.2 Second section results

ROC curves of the algorithms applied to 3D PCD in both, original and alternate data sets are presented in Figures 4.79 and 4.80, respectively.

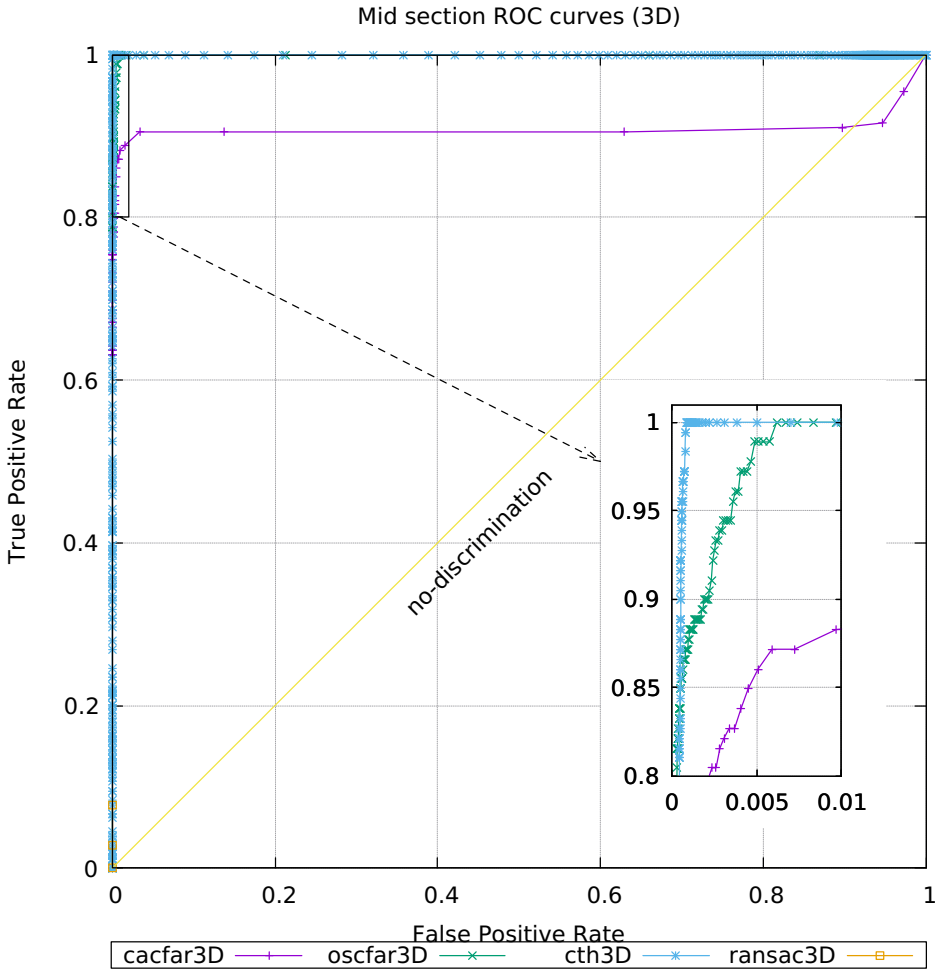

FIGURE 4.79: ROC curves of the methods applied to 3D PCD in the mid section of underground mine tunnel data set.

Segmentation threshold, as expected, presents a high level of performance, closely followed by OS-CFAR 3D. As in the previous section's results, segmentation threshold is affected by the artificially incorporated clutter in the alternate data set. CA-CFAR methods, on the other hand, are not affected by the clutter in the alternate data set.

ROC curves corresponding to 2D projected data for the original and alternate data set are shown in Figures 4.81, 4.82, 4.83 and 4.84.

It is not possible to observe a significant difference between the ROC curves applied to equirectangular projected data and to Mercator projected data. Also, the the results in the

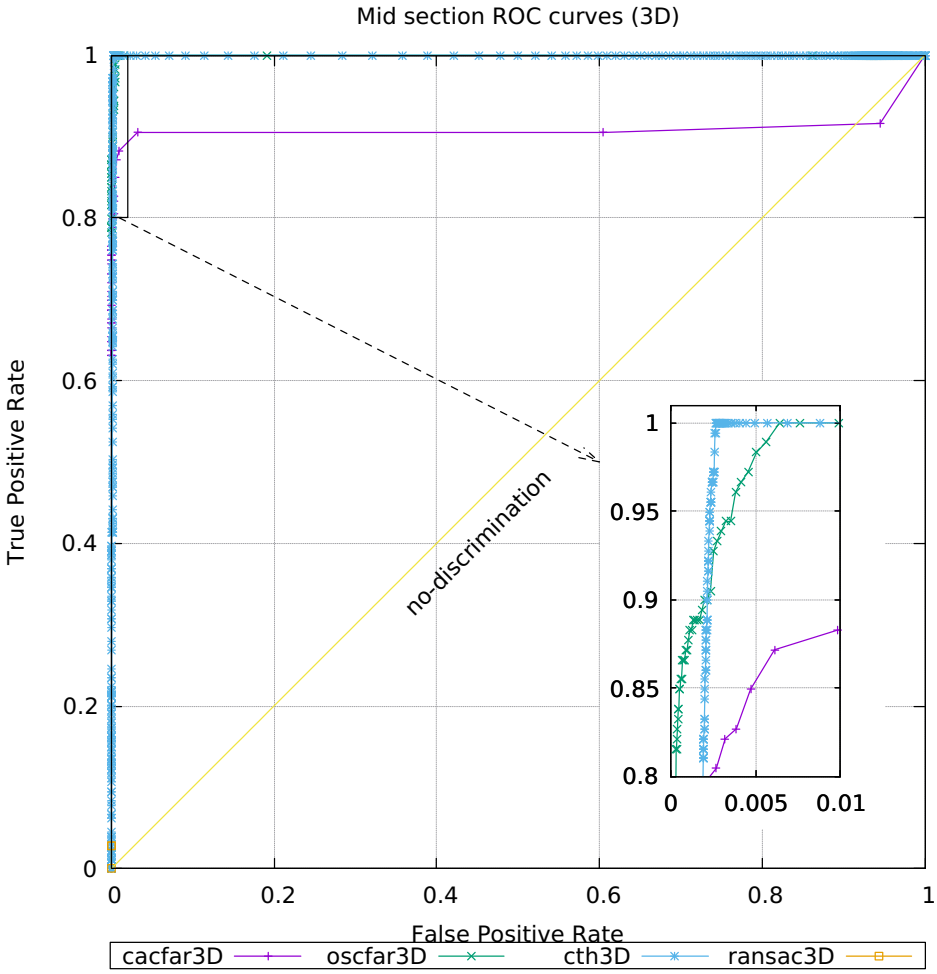

FIGURE 4.80: ROC curves of the methods applied to 3D PCD in the mid section of underground mine tunnel data set with artificial clutter.

original and alternate data set are comparable.

The methods applied to 2D projected data are ranked in table 4.12 according to their performance in the ROC curves.

In this data set, the standard versions of the CFAR methods exhibit the best results, followed by segmentation thresholding. Adaptive versions of the CFAR methods are ranked after segmentation thresholding. In all cases, the OS-CFAR versions present a better performance than the CA-CFAR counterpart.

Detection probability and actual false alarm probability against required false alarm rate are presented in Figures 4.85, 4.86, 4.87 and 4.88 for both, original and alternate data sets, respectively, in the same fashion as in the previous section.

OS-CFAR 3D presents highest detection rates in both, the original and the alternate data set. CA-CFAR 3D also show higher detection rates than all the 2D CFAR methods. In the

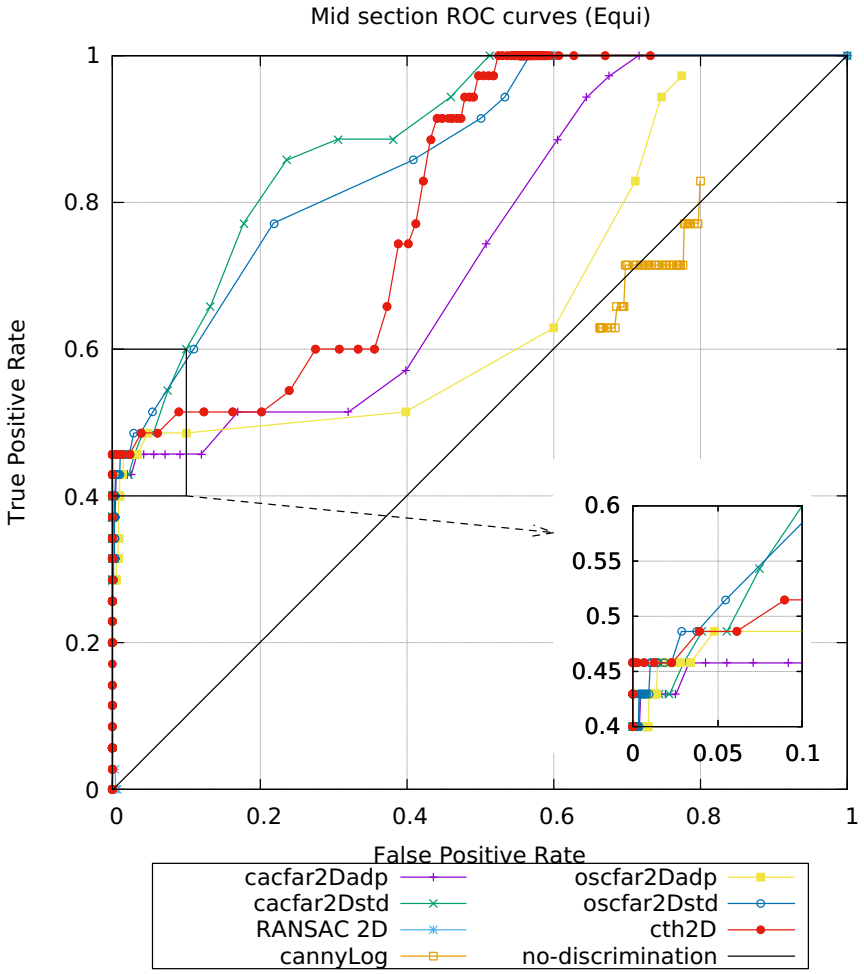

FIGURE 4.81: ROC curves of the methods applied to equirectangular 2D projected data in the mid section of underground mine tunnel data set.

original data set all the CFAR methods applied to 2D projected data present comparable detection rates, while in the alternate data set OS-CFAR adaptive methods show higher detection rates than the other 2D CFAR algorithms.

On the other hand, CA-CFAR standard versions exhibit the best performance with respect to the actual false alarm rate in the original data set. However, all methods present a lower performance in terms of actual false alarm rate in the alternate data set.

Table 4.13 ranks all the CFAR methods according to their results in in Figures 4.85, 4.86, 4.87 and 4.88.

In the alternate data set, CA-CFAR 3D presents good performance in therms of both, detection rate and actual false alarm rate. Also, in the original data set, CA-CFAR 3D exhibits a similar performance for required false alarm rate values greater than 1×10^{-3}. In terms of the actual false alarm rate, CA-CFAR methods are ranked higher than the OS-CFAR

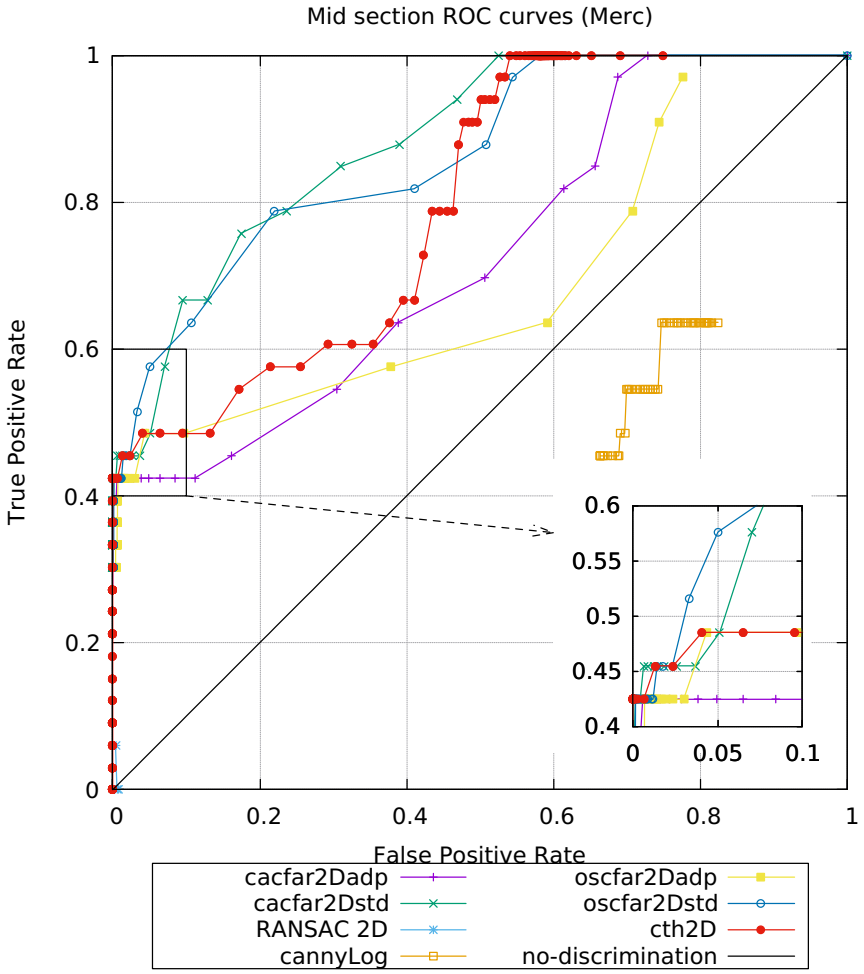

FIGURE 4.82: ROC curves of the methods applied to Mercator 2D projected data in the mid section of underground mine tunnel data set.

algorithms.

The OSPA metric results are shown in Figures 4.89, 4.90 and 4.91 (original clutter version), as well as, in Figures 4.92, 4.93 and 4.94 (high clutter version).

OSPA errors in 2D projected data for TPR values higher than 0.5 are dominated by the high number of false alarm, for all methods. For TPR valuse lower than 0.5, segmentation threshold yields lower OSPA errors, due to the high SNR, as discussed in the results of the other section of the mine. In the dataset version with original clutter, CA-CFAR adaptive and standard versions present OSPA distances comparable to the segmentation threshold OSPA performance. On the other hand, all methods are affected by the artificially introduced clutter, presenting higher OSPA distances.

In the case of 3D PCD data, CA-CFAR 3D presents lower OSPA errors with TPR values ranging between 0.6 and 0.75, in both, original and artificial clutter dataset versions. Con-

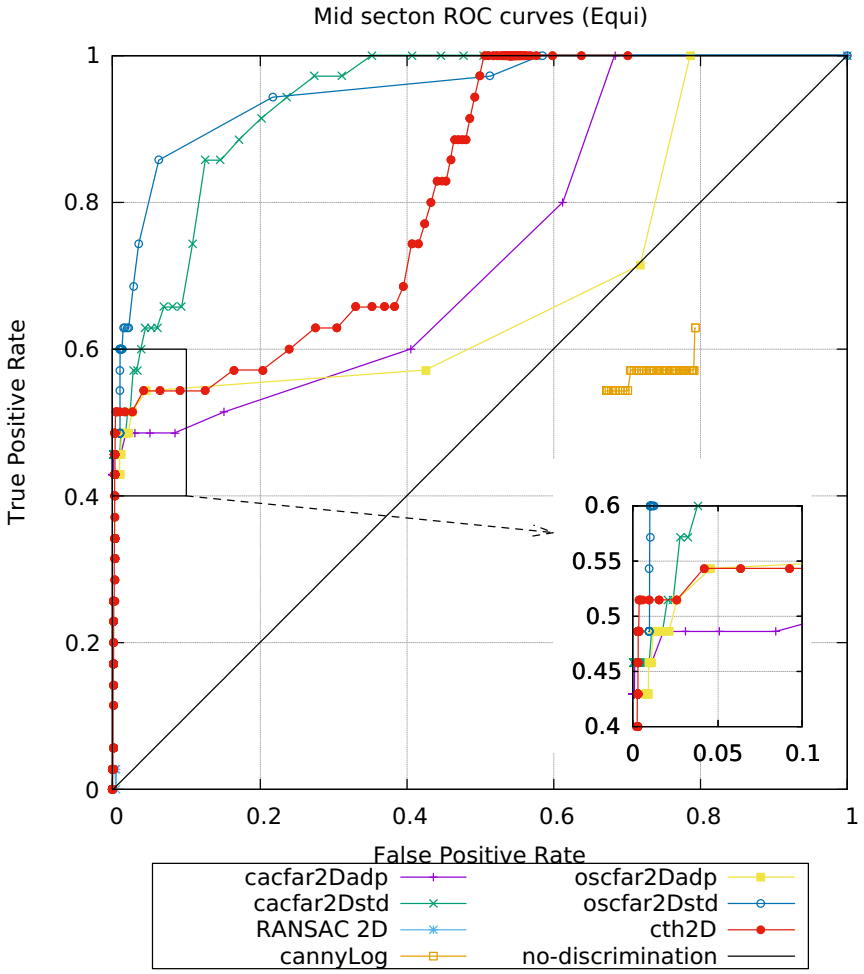

FIGURE 4.83: ROC curves of the methods applied to equirectangular 2D projected data in the mid section of underground mine tunnel data set (artificial clutter).

sistent with the ROC results above, OS-CFAR 3D achieves an OSPA distance lower than segmentation threshold with a TPR value around 0.85, in the high clutter dataset version.

It was observed in all datasets that the OSPA metric was heavily affected by the output of the CCA method, particularly in the experimental datasets. Direct application of the OSPA metric to the points detected by the methods (without using the CCA algorithm) was not possible because of the high number of detections and the high computational complexity and high computer memory requirements of the OSPA algorithm.

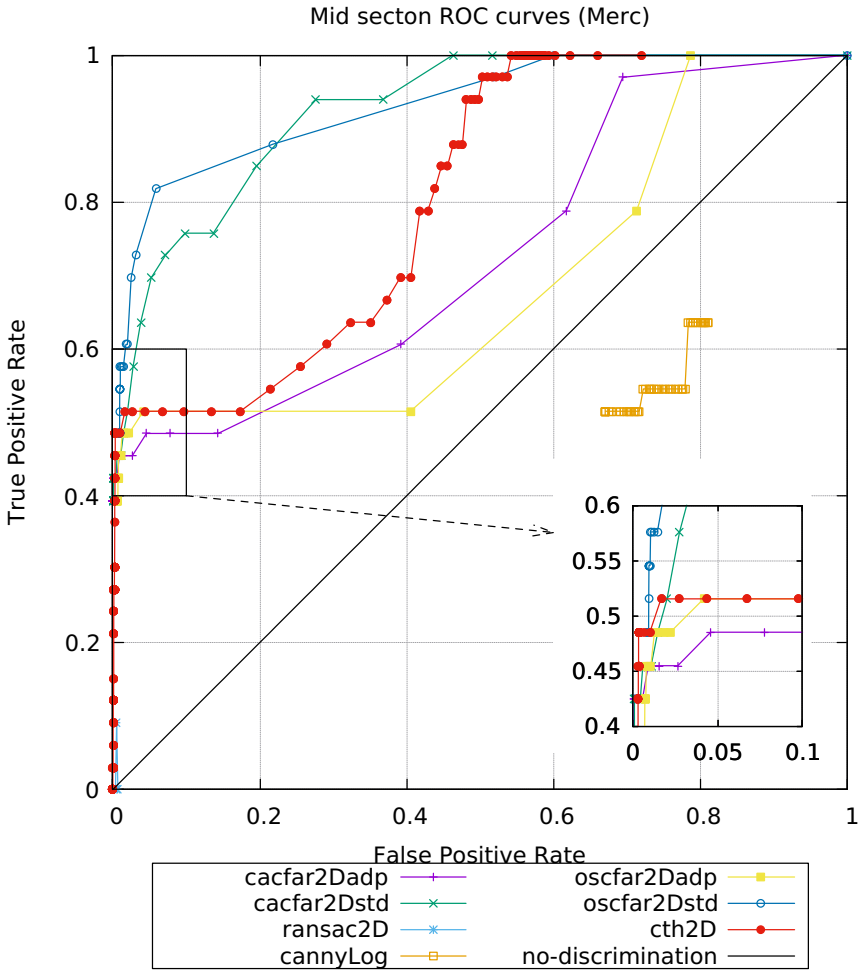

FIGURE 4.84: ROC curves of the methods applied to Mercator 2D projected data in the mid section of underground mine tunnel data set (artificial clutter).

TABLE 4.12: Ranking of 2D methods based on ROC curves for FPR < 0.1, "Original" refers to the original data set without additional clutter, while the "Alternate" corresponds to a copy of the data set with patch of added clutter points.

Rank	Original (Equi)	Original (Merc)	Alternate (Equi)	Alternate (Merc)
1	oscfar2Dstd	oscfar2Dstd	oscfar2Dstd	oscfar2Dstd
2	cacfar2Dstd	cacfar2Dstd	cacfar2Dstd	cacfar2Dstd
3	cth2D	cth2D	cth2D	cth2D
4	oscfar2Dadp	oscfar2Dadp	oscfar2Dadp	oscfar2Dadp
5	cacfar2Dadp	cacfar2Dadp	cacfar2Dadp	cacfar2Dadp
6	RANSAC 2D	RANSAC 2D	RANSAC 2D	RANSAC 2D
7	cannyLog	cannyLog	cannyLog	cannyLog

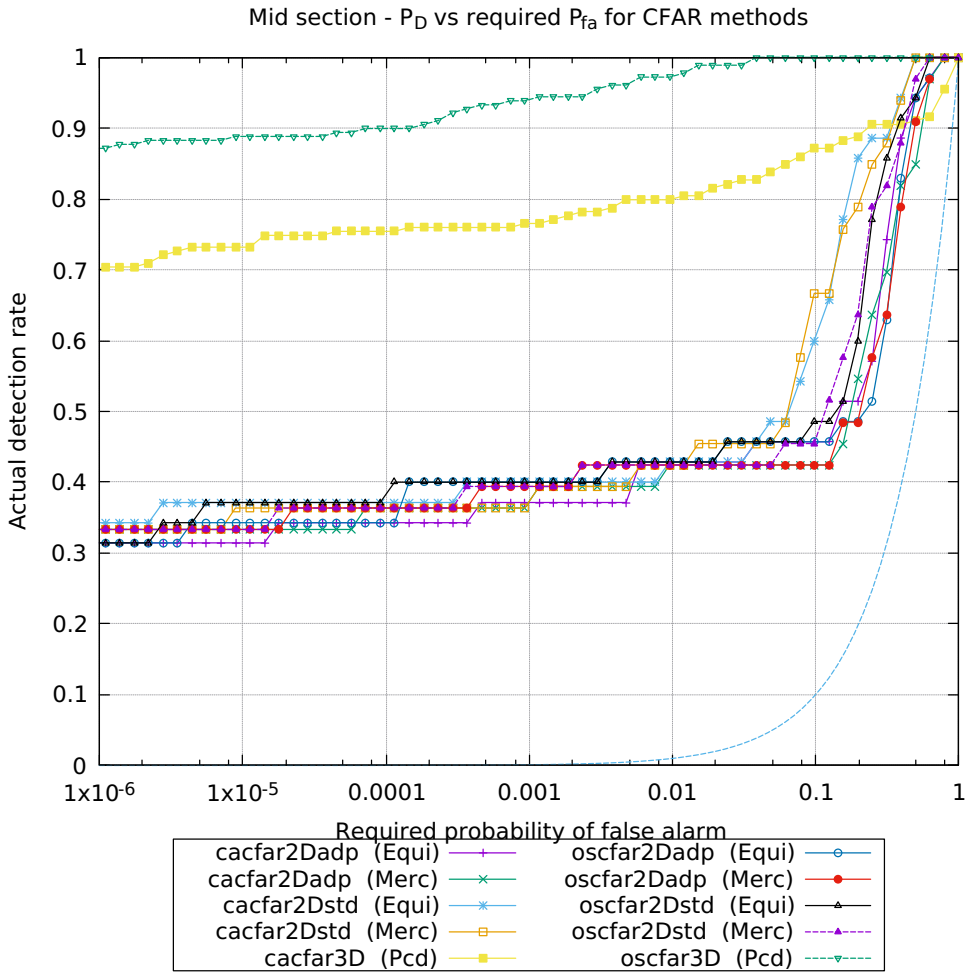

FIGURE 4.85: Probability of detection vs. required false alarm rate for mid section of underground mine tunnel data.

TABLE 4.13: Ranking of CFAR methods based on probability of detection (TPR) and actual false alarm rate (FPR).

Rank	Original TPR	Original actual FPR	Alternate TPR	Alternate actual FPR
1	oscfar3D (Pcd)	cacfar2Dstd (Merc)	oscfar3D (Pcd)	cacfar3D (Pcd)
2	cacfar3D (Pcd)	cacfar2Dstd (Equi)	cacfar3D (Pcd)	cacfar2Dadp (Merc)
3	cacfar2Dstd (Equi)	cacfar2Dadp (Merc)	oscfar2Dstd (Equi)	cacfar2Dadp (Equi)
4	oscfar2Dstd (Equi)	cacfar2Dadp (Equi)	oscfar2Dstd (Merc)	cacfar2Dstd (Equi)
5	oscfar2Dadp (Equi)	cacfar3D (Pcd)	cacfar2Dstd (Equi)	cacfar2Dstd (Merc)
6	oscfar2Dadp (Merc)	oscfar2Dstd (Merc)	oscfar2Dadp (Equi)	oscfar3D (Pcd)
7	oscfar2Dstd (Merc)	oscfar3D (Pcd)	cacfar2Dadp (Equi)	oscfar2Dadp (Merc)
8	cacfar2Dstd (Merc)	oscfar2Dstd (Equi)	oscfar2Dadp (Merc)	oscfar2Dadp (Equi)
9	cacfar2Dadp (Merc)	oscfar2Dadp (Merc)	cacfar2Dstd (Merc)	oscfar2Dstd (Merc)
10	cacfar2Dadp (Equi)	oscfar2Dadp (Equi)	cacfar2Dadp (Merc)	oscfar2Dstd (Equi)

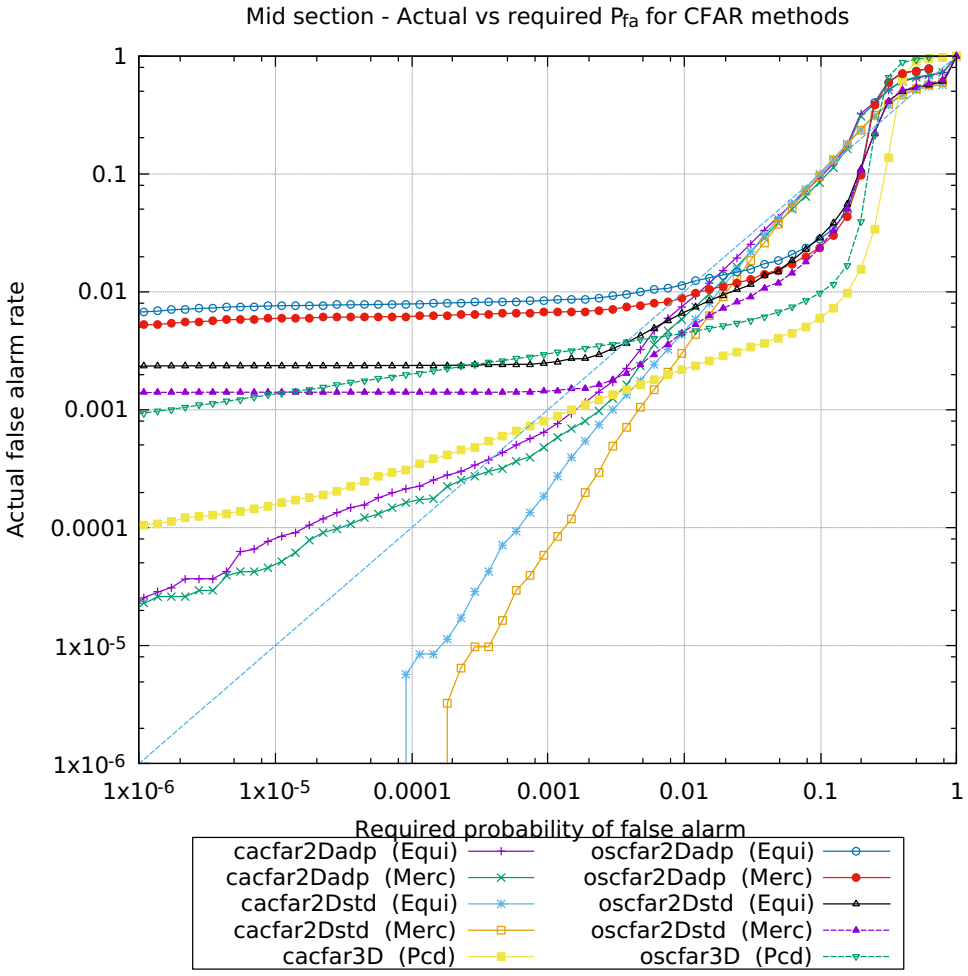

FIGURE 4.86: Probability of false alarm vs. required false alarm rate for mid section of underground mine tunnel data.

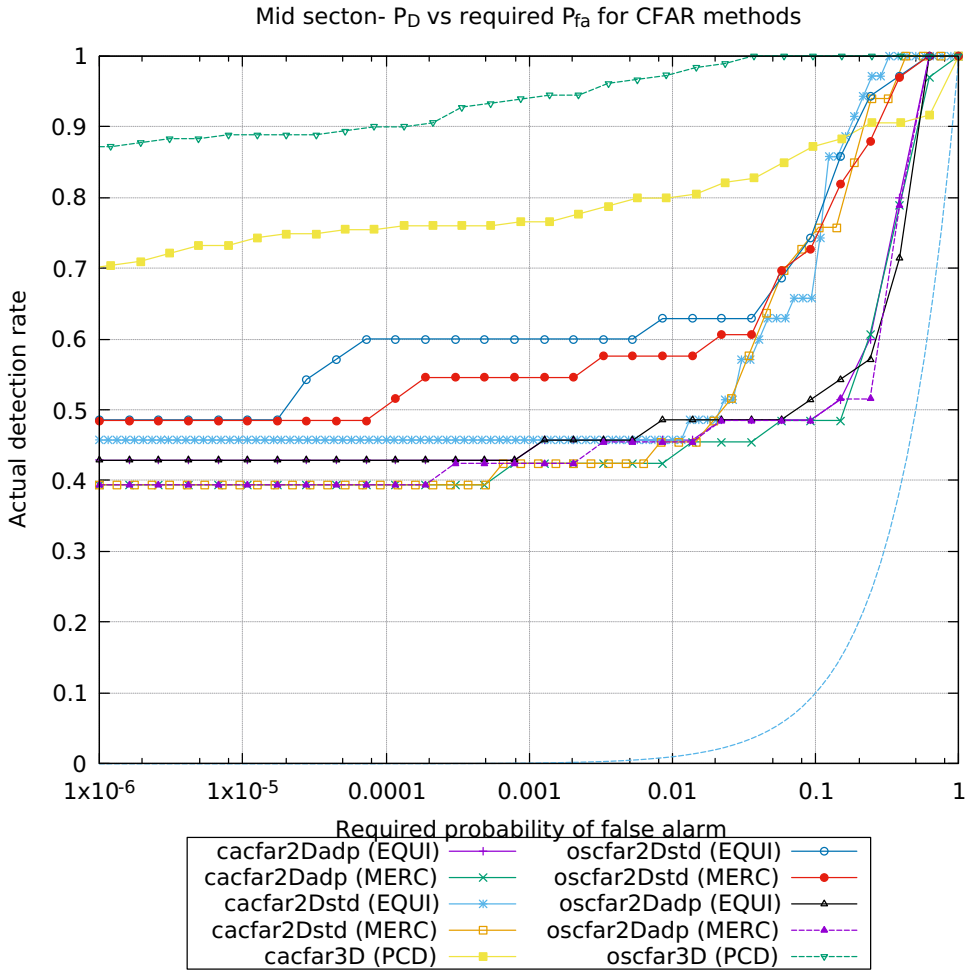

FIGURE 4.87: Probability of detection vs. required false alarm rate for mid section of underground mine tunnel data (artificial clutter).

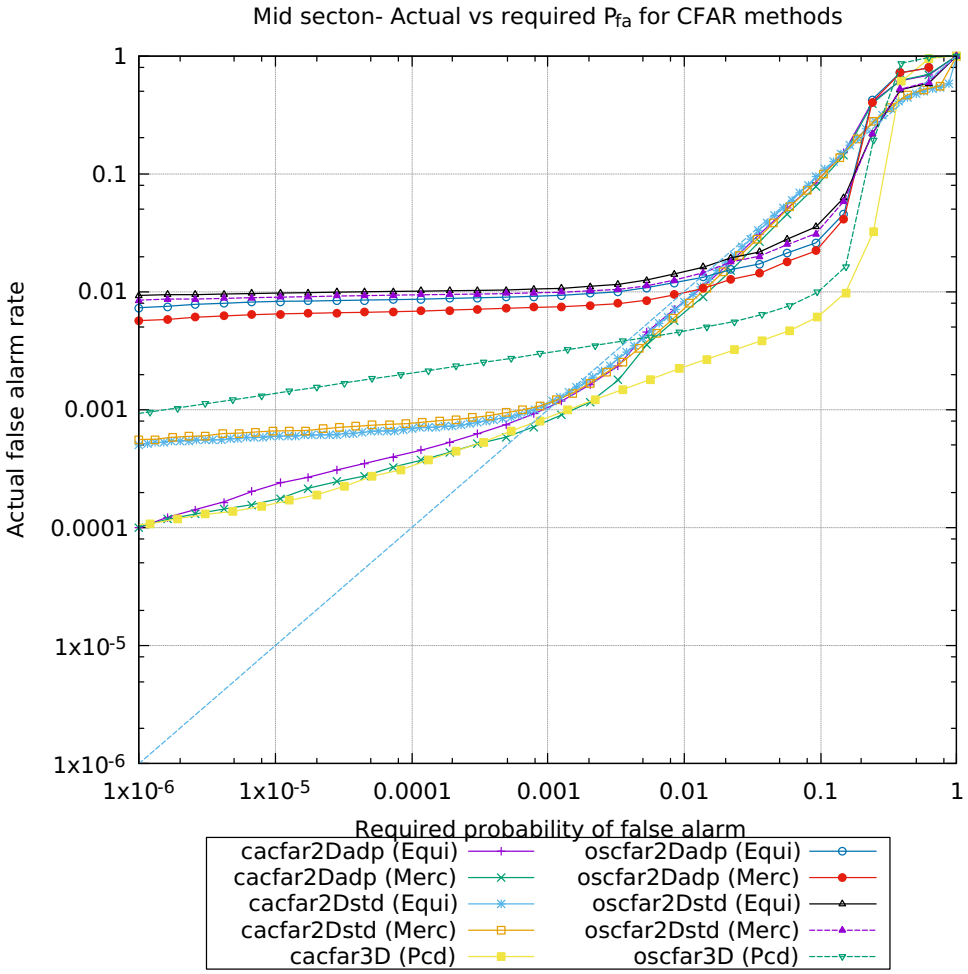

FIGURE 4.88: Probability of false alarm vs. required false alarm rate for mid section of underground mine tunnel data (artificial clutter).

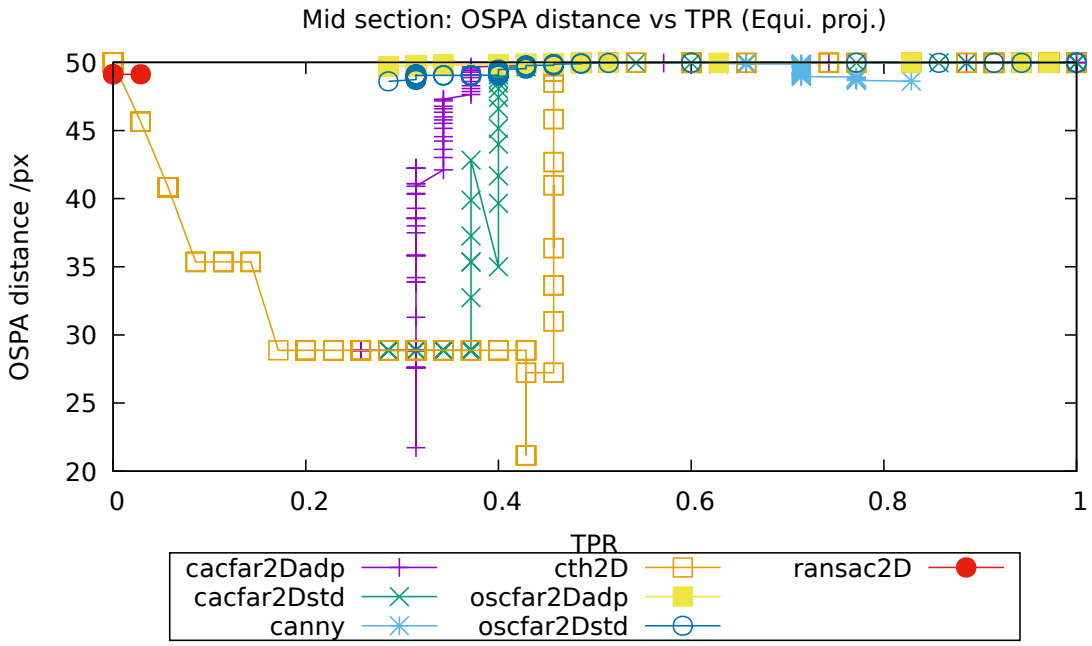

FIGURE 4.89: OSPA metric results for the mid section of undergroun mine tunnel (Equirectangular projection).

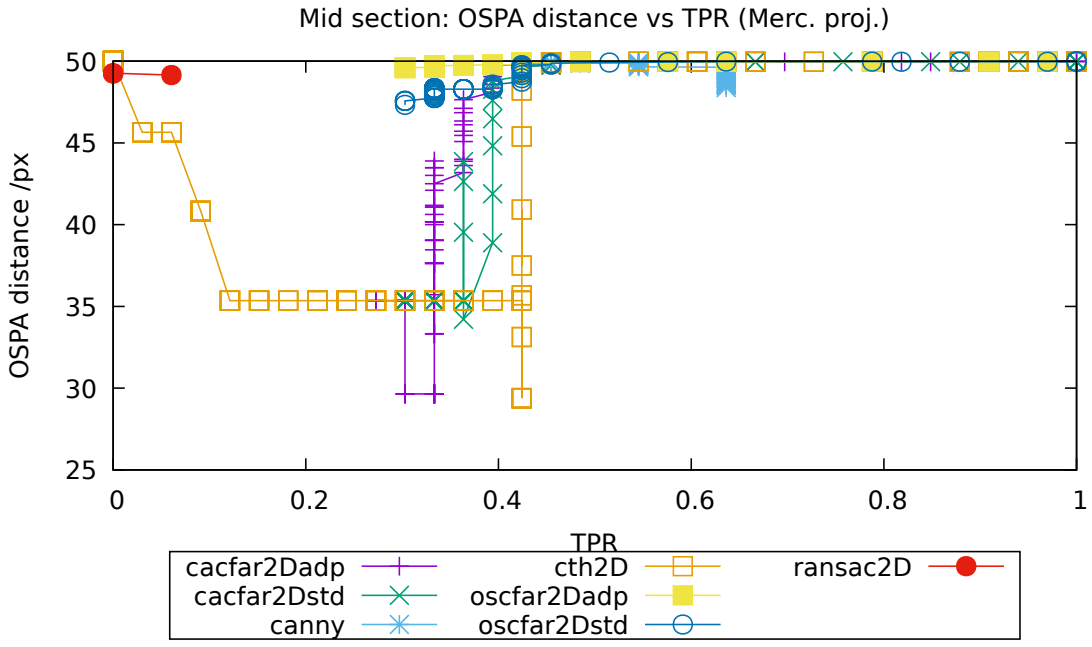

FIGURE 4.90: OSPA metric results for the mid section of undergroun mine tunnel (Mercator projection).

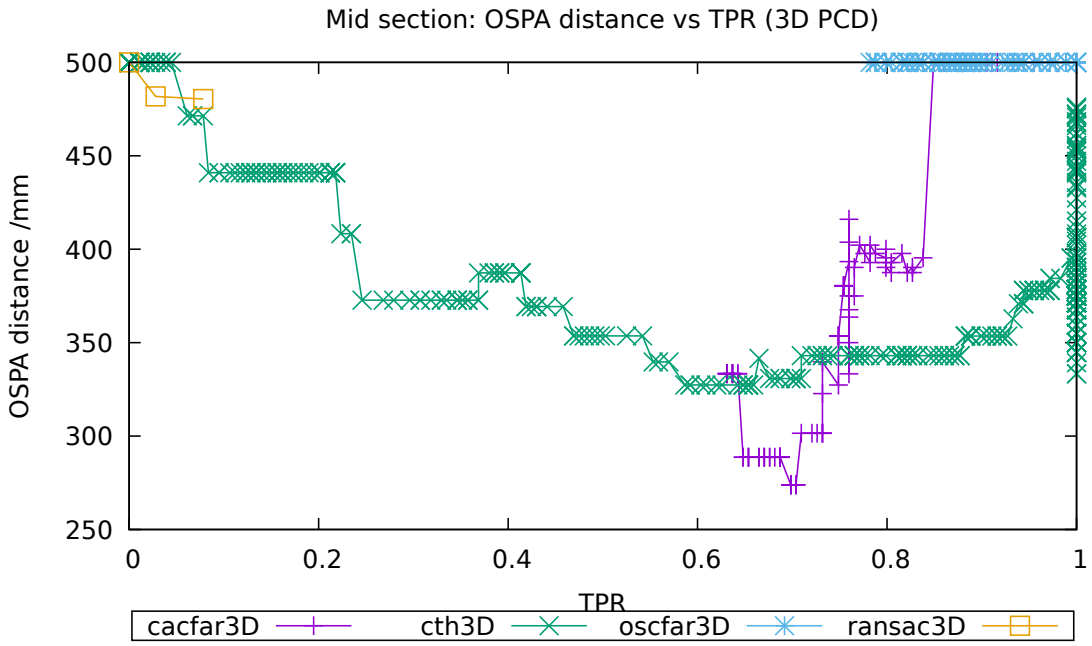

FIGURE 4.91: OSPA metric results for the mid section of undergroun mine tunnel (3D PCD).

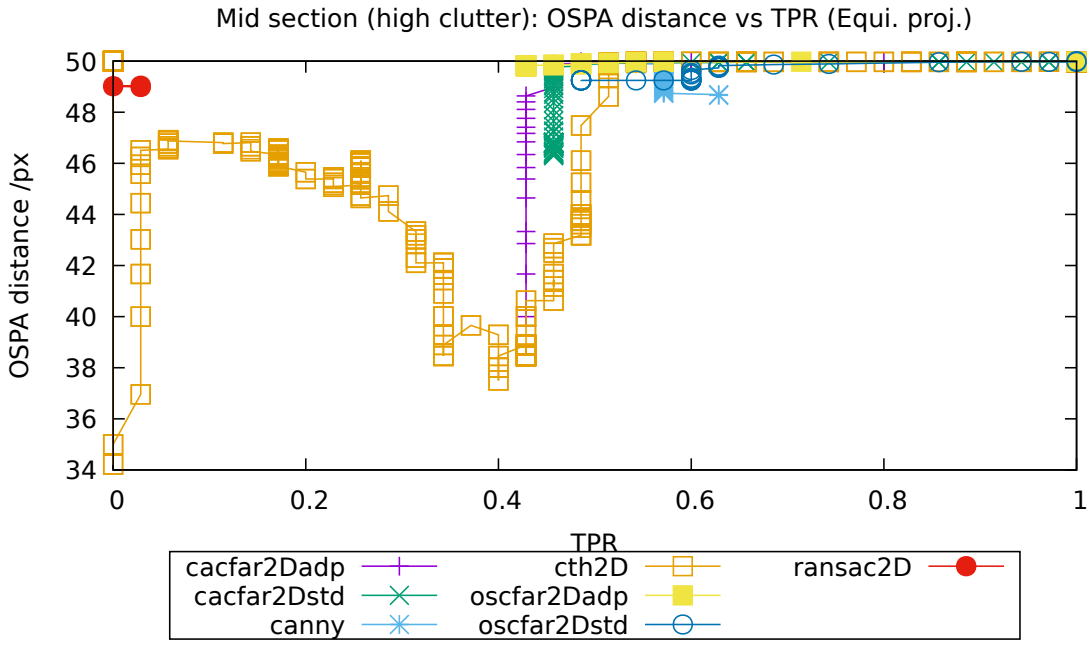

FIGURE 4.92: OSPA metric results for the mid section of undergroun mine tunnel with high clutter (Equirectangular projection).

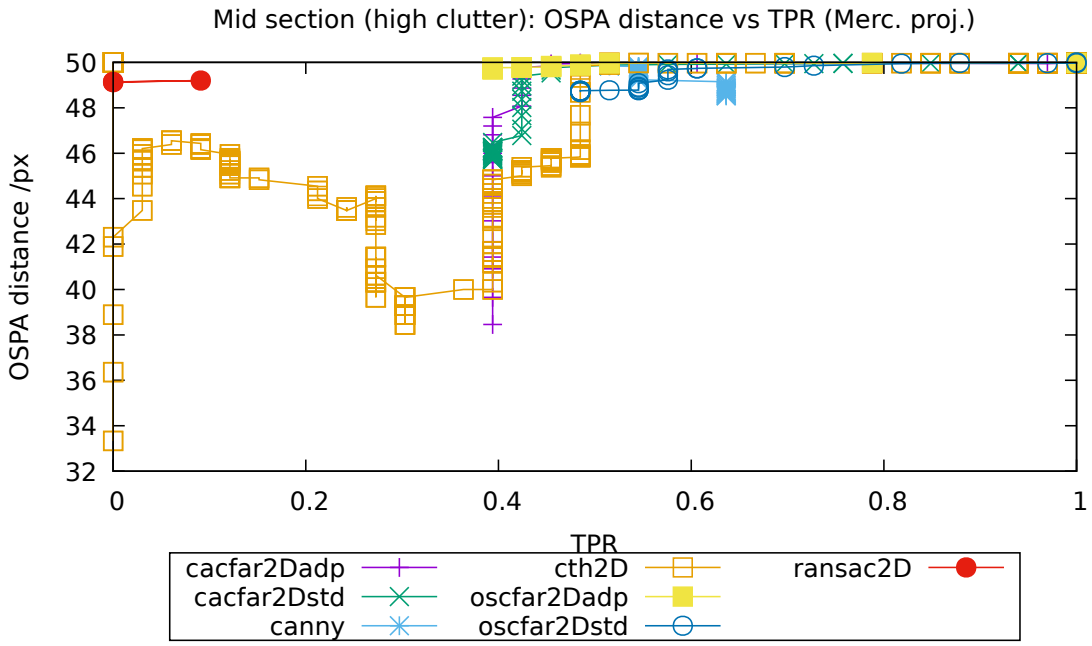

FIGURE 4.93: OSPA metric results for the mid section of undergroun mine tunnel with high clutter (Mercator projection).

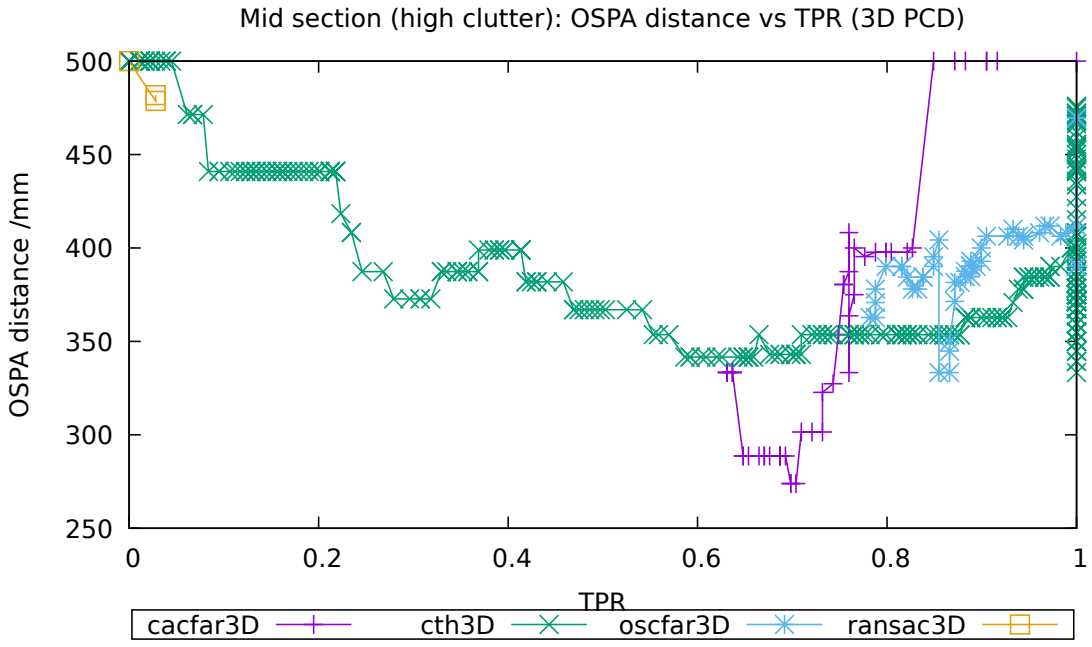

FIGURE 4.94: OSPA metric results for the mid section of undergroun mine tunnel with high clutter (3D PCD).

4.9 Overall discussion on the CFAR methods' performance

In situations of high SNR which are optimal for Segmentation Thresholding, the CFAR methods yielded comparable performance. In all situations with a certain level of clutter, the CFAR methods achieved a higher performance than Segmentation Thresholding. The Canny Edge Detector is very sensitive to noise, which was observed in most situations with simulated or real data. The RANSAC method was also sensitive to noise in its 2D version, while the spatial data structure in 3D also affected its performance (several different structures in the 3D PCD were consistent with the cylinder model, and therefore detected as false alarms). Nevertheless, this algorithm can be used after the detection phase to check whether the detected objects match a particular shape and size.

In most cases the 3D CFAR version exhibits a better performance than the 2D CFAR versions. The main disadvantage of 3D CFAR methods arises with Point Clouds with a large number of points, due to the computationally expensive process of building the KD-Tree and iterating over all points calculating the local statistics. The KD-Tree is needed to efficiently obtain the points close to the cell under test which correspond to reference and guard cells. This effect can be minimised if the Point Cloud is partitioned to only include volumes where targets are known to be present. If that is not possible, and faster calculations are required, then projecting the data into 2D, using a simple transformation such as the Equirectangular projection, may be advisable. When working with 2D data and it is known that the targets have very different apparent sizes, the CFAR adaptive window size versions should be used. Otherwise, when the targets' apparent size does not vary significantly, it is better to use the standard versions, as the adaptive version requires more time. As a general remark, the adaptive OS-CFAR is much slower than the adaptive CA-CFAR, due to the non-linear solver that needs to be executed every time the reference cell window size changes. The OS-CFAR processor should be used when it is expected that part of the points representing the target will spread into the reference cell window (i.e. when detecting large objects) or when it is known that the noise distribution does not match the one expected by the CA-CFAR detector. OS-CFAR methods presented, in general, a better performance than the CA-CFAR methods, particularly in the experimental data sets. Nevertheless, in many situations the difference in performance is not large, and the fastest CA-CFAR approach can be considered. In summary, the results confirm that the adaptive window size version of the 2D CFAR processor yields a better performance when the targets are located a different distances from the sensor. This is due to the adaptive CFAR window which grows or shrinks based on the real size of the object (in mm) and not a fixed size in pixels (or bins) which is

the standard case. The 3D version of CFAR processor also showed advantages over the 2D version of the methods because there is no distortion in the objects size or shape.

Chapter 5

Conclusions

5.1 Final analysis and remarks

In Chapter 3 a noise reduction method based on the Binary integrator detector and spectral noise subtraction has been introduced. Its performance has been compared with two other general, widely used, noise reduction methods, Wiener filtering and wavelet denoising.

The BINR method exhibits good noise reduction capabilities, which result in a higher signal-to-noise power ratio than the other two techniques. Importantly, BINR has the ability to preserve most of the signal's spatial details. Also, the BINR algorithm's computational time is lower than the other two methods. The reduced time complexity shown by the BINR method and its denoising capabilities make it an appealing noise reduction alternative to be used as a pre-processing step for radar data.

The BINR method requires multiple radar data observations of the scene. This is due to the Binary Integration detector on which the method is based. Integration is a widely used technique in different scientific fields and technological applications to reduce noise by increasing the signal to noise ratio. It is explained by the Central Limit Theorem from probability theory. By taking more samples of the phenomenon under study, the sampling distribution of the mean of samples approaches a normal distribution and its standard deviation decreases with the squared root of the number of samples. Therefore, the sensor should remain in a static location for a certain amount of time, before moving to a new location, in a "stop-scan-move" strategy. This strategy is commonly used in robot exploration missions or when high accuracy is required. Alternatively, if a registration algorithm is applied to align multiple observations from different scanning positions into a common reference frame then, the overlapping regions can also be processed with the BINR algorithm. On the other hand, the BINR method has been designed to be used as an *on-line* technique, i.e. it can process

data as it is captured by the sensor.

Three approaches to detect reflective markers in 3D ladar surveying data, based on CA-CFAR processors have been introduced in Chapter 4. Their performances have been compared with a standard, constant threshold, segmentation method from image processing. The first method consists in the direct application of a standard 2D CA-CFAR processor on the 2D projection of the data. The second technique introduces a new adaptive window 2D CA-CFAR processor also applied to the data projected in 2D. This method combines intensity and range information. The latter is used to adapt the window size parameter on each cell under test. The third approach corresponds to a 3D implementation of the CA-CFAR algorithm to be used directly on the 3D ladar data.

Two simulated environments and two experimental data sets were used to test these methods, an indoor laboratory data set and an underground mine tunnel. In conditions of high noise and clutter, the CFAR methods present higher detection and/or lower false alarm rates than other state-of-the-art methods used for comparison.

5.2 Future work

The encouraging results obtained with the BINR method open the door for several improvements for radar data noise reduction. One track for improvement lies in enhancing the standard Wiener or Wavelet methods by making them adaptable to different noise distributions. This approach has been taken for some wavelet denoising applications in which different threshold functions are derived for some non-Gaussian noise distributions [CH05, ZC00, ALP02]. This can then be the basis for the integration of 2 or more of the methods. This has been carried out for Wiener and Wavelet filtering in the work of Jin et al. [JFWJ03]. This work exploited the advantages of both methods simultaneously, but even though the theoretical improvement in peak-to-peak SNR was expected to be 3dB, only 0.5dB was achieved. Also the time complexity of Jin et al's method corresponds to the combination of both the Wiener and wavelet methods. In particular, it would be interesting to evaluate optimized combinations of the Wiener filter and BINR, Wavelet denoising and BINR, and the Wiener-Wavelet and the BINR method, in terms of their computational complexities. These combinations would be expected to yield the advantages of the individual methods, for instance, the smooth background noise presented by the Wiener filter, the reduction in the mean noise power value provided by the BINR method, and the faithful representation of the original signal achieved when applying wavelet denoising.

One route for improvements for the methods developed in Chapter 4 consists in extending other CFAR methods, which would allow their application to a broader range of use cases,

other than the 3D lidar surveying case studied in this work. Also, the 2D versions could be extended to more 2D projections, which might be better suited in other applications. Both, the 3D and the 2D adaptive window CFAR methods could benefit from memory usage and time complexity optimisations, particularly for large data sets. Following the 2D to 3D extension of the CFAR processors, it might be possible to generalise the algorithms to N-dimensional data sets. Another, important, line of future work consists in testing the results from the detectors presented in this work as starting estimates in a full registration scheme, such as the Iterative Closest Points (ICP), described in Section 4.2.1.

Chapter 6

Bibliography

[AJ12] M. Adams and E. Jose. *Robotic Navigation and Mapping with Radar*. Artech House, 2012.

[ALP02] A. Antoniadis, D. Leporini, and J. C. Pesquet. Wavelet thresholding for some classes of non-Gaussian noise. *Statistica Neerlandica*, 56(4):434–453, 2002.

[AOE06] O. A. M. Aly, A. S. Omar, and A. Z. Elsherbeni. Detection and localization of rf radar pulses in noise environments using wavelet packet transform and higher order statistics. *Progress In Electromagnetics Research*, 58:301–317, 2006.

[BAL+17] P. Barrios, M. Adams, K. Leung, F. Inostroza, G. Naqvi, and M. E. Orchard. Metrics for evaluating feature-based mapping performance. *IEEE Transactions on Robotics*, 33(1):198–213, Feb 2017.

[Bar05] M. Barkat. *Signal detection and estimation*. Artech House radar library. Artech House, 2005.

[BB07] J. Böhm and S. Becker. Automatic marker-free registration of terrestrial laser scans using reflectance features. In *Proceedings of 8th Conference on Optical 3D Measurement Techniques, Zurich, Switzerland*, pages 338–344, 2007.

[BDR08] C Brenner, C Dold, and N Ripperda. Coarse orientation of terrestrial laser scans in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(1):4–18, 2008.

[Ben75] Jon Louis Bentley. Multidimensional binary search trees used for associative searching. *Commun. ACM*, 18(9):509–517, September 1975.

[BETVG08] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). *Computer vision and image understanding*, 110(3):346–359, 2008.

[BF08] Shahar Barnea and Sagi Filin. Keypoint based autonomous registration of terrestrial laser point-clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(1):19–35, 2008.

[BGG98] C.S. Burrus, R.A. Gopinath, and H. Guo. *Introduction to wavelets and wavelet transforms: a primer*, volume 23. Prentice hall Upper Saddle River NJ, 1998.

[BGG+04] P. Benz, M. Geser, R. Graf, B. Pfiffner, M. Buerki, M. Landert, M. Hammerer, J. Hinderling, G. Hanns, and G. Bayer. Process and device for the automatic location of reference markers, May 11 2004. US Patent 6,734,952.

[BKE+95] J.C. Brailean, R.P. Kleihorst, S. Efstratiadis, A.K. Katsaggelos, and R.L. Lagendijk. Noise reduction filters for dynamic image sequences: a review. *Proceedings of the IEEE*, 83(9):1272 –1292, sep 1995.

[BM92a] P. J. Besl and H. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, Feb 1992.

[BM92b] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Robotics-DL tentative*, pages 586–606. International Society for Optics and Photonics, 1992.

[Bol79] S. Boll. Suppression of acoustic noise in speech using spectral subtraction. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 27(2):113 – 120, apr 1979.

[BPR12] Sandro Barone, Alessandro Paoli, and Armando Viviano Razionale. Three-dimensional point cloud alignment detecting fiducial markers by structured light stereo imaging. *Machine Vision and Applications*, 23(2):217–229, 2012.

[Bro92] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM Comput. Surv.*, 24(4):325–376, December 1992.

[Bru90] J.S. Brunner. A recursive method for calculating binary integration detection probabilities. *IEEE Transactions on Aerospace and Electronic Systems*, 26(6):1034–1035, November 1990.

[BSLK04] Yaakov Bar-Shalom, X Rong Li, and Thiagalingam Kirubarajan. *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.

[BSM79] M. Berouti, R. Schwartz, and J. Makhoul. Enhancement of speech corrupted by acoustic noise. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '79.*, volume 4, pages 208 – 211, apr 1979.

[BSSC95] A.S.E. Board, C.E.T. Systems, D.E.P. Sciences, and N.R. Council. *The Global Positioning System: A Shared National Asset*. Romantic Reassessment; 118. National Academies Press, 1995.

[Can86] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, Nov 1986.

[CC96] Mu-Yen Chen and Jung-Jae Chao. Radar image denoising by recursive thresholding. In *Image Processing, 1996. Proceedings., International Conference on*, volume 1, pages 395 –398 vol.1, sep 1996.

[CH05] Y. Chen and C. Han. Adaptive wavelet threshold for image denoising. *Electronics Letters*, 41(10):586 – 587, may 2005.

[CMCTN07] José A Castellanos, Ruben Martinez-Cantin, Juan D Tardós, and José Neira. Robocentric map joining: Improving the consistency of ekf-slam. *Robotics and autonomous systems*, 55(1):21–29, 2007.

[dBG01] M. di Bisceglie and C. Galdi. Cfar detection of extended objects in high resolution sar images. In *Geoscience and Remote Sensing Symposium, 2001. IGARSS '01. IEEE 2001 International*, volume 6, pages 2674–2676 vol.6, 2001.

[DJ94a] David L. Donoho and Iain M. Johnstone. Ideal denoising in an orthonormal basis chosen from a library of bases. *Comptes Rendus Acad. Sci., Ser. I*, 319:1317–1322, 1994.

[DJ94b] David L. Donoho and Jain M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[DJ94c] D.L. Donoho and I.M. Johnstone. Threshold selection for wavelet shrinkage of noisy data. In *Engineering in Medicine and Biology Society, 1994. Engineering Advances: New Opportunities for Biomedical Engineers. Proceedings of the 16th Annual International Conference of the IEEE*, pages A24 –A25 vol.1, nov 1994.

[Don95] D.L. Donoho. De-noising by soft-thresholding. *Information Theory, IEEE Transactions on*, 41(3):613 –627, may 1995.

[DT12] Ioannis N Daliakopoulos and Ioannis K Tsanis. A weather radar data processing module for storm analysis. *Journal of hydroinformatics*, 14(2):332–344, 2012.

[EHS+14] F. Endres, J. Hess, J. Sturm, D. Cremers, and W. Burgard. 3-d mapping with an rgb-d camera. *IEEE Transactions on Robotics*, 30(1):177–187, Feb 2014.

[Ein12] Garry Einicke. *Smoothing, Filtering and Prediction - Estimating The Past, Present and Future*, chapter Discrete-Time Minimum-Mean-Square-Error Filtering, pages 25 – 40. InTech, 2012.

[FB81] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, June 1981.

[FBW01a] A. Foessel, J. Bares, and W.R.L. Whittaker. Three-Dimensional Map Building with MMW RADAR. In *Proceedings of the 3rd International Conference on Field and Service Robotics*, Helsinki, Finland, June 2001.

[FBW01b] Alex Foessel, John Bares, and William Red L. Whittaker. Three-dimensional map building with MMW radar. In *Proceedings of the 3rd International Conference on Field and Service Robotics*, Helsinki, Finland, June 2001.

[FCA99] Alex Foessel, Sachin Chheda, and Dimitrios Apostolopoulos. Short-range millimeter-wave radar perception in a polar environment. In *Proceedings of the Field and Service Robotics Conference*, August 1999.

[Fia10] M. Fiala. Designing highly reliable fiducial markers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(7):1317–1324, July 2010.

[Foe02a] A. Foessel. *Scene Modeling from Motion-Free Radar Sensing*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, January 2002.

[Foe02b] Alex Foessel. *Scene Modeling from Motion-Free Radar Sensing*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, January 2002.

[Fra04] András Frank. On kuhn's hungarian method – a tribute from hungary. Technical report, Eötvös Loránd University, 2004.

[GK94] P.P. Gandhi and S.A. Kassam. Optimality of the cell averaging CFAR detector. In *IEEE Transactions on Information Theory*, volume 40, pages 1226–1228, July 1994.

[GLZ+09] Gui Gao, Li Liu, Lingjun Zhao, Gongtao Shi, and Gangyao Kuang. An adaptive and fast cfar algorithm based on automatic censoring for target detection in high-resolution sar images. *Geoscience and Remote Sensing, IEEE Transactions on*, 47(6):1685–1697, June 2009.

[Gos05] A. Ardeshir Goshtasby. *2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications*. Wiley, 2005.

[HCBD07a] Peter Hansen, Peter Corke, Wageeh Boles, and Kostas Daniilidis. Scale invariant feature matching with wide angle images. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 1689–1694. IEEE, 2007.

[HCBD07b] Peter Hansen, Peter Corke, Wageeh Boles, and Kostas Daniilidis. Scale-invariant features on the sphere. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.

[HEBN13] H Houshiar, Jan Elseberg, Dorit Borrmann, and Andreas Nüchter. A study of projections for key point based registration of panoramic terrestrial 3d laser scans. *Geo-spatial Information Science (GSIS)*, 2013.

[HHD10] Gibson Hu, Shoudong Huang, and Gamini Dissanayake. Evaluation of pose only slam. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 3732–3737. IEEE, 2010.

[Hub02] Daniel F Huber. *Automatic three-dimensional modeling from reality*. PhD thesis, Carnegie Mellon University, 2002.

[JA05] E. Jose and M.D. Adams. An augmented state SLAM formulation for multiple line-of-sight features with millimetre wave radar. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Alberta, Canada, August 2005.

[JAMP10] E. Jose, M.D. Adams, J. Mullane, and N. Patrikalakis. Predicting millimetre wave radar spectra for autonomous navigation. *IEEE Sensors Journal*, 10(5):960–971, May 2010.

[JB08] D. Johnson and G. Brooker. Research radar for unmanned navigation. In *Radar, 2008 International Conference on*, pages 165 –170, sept. 2008.

[JFWJ03] F. Jin, P. Fieguth, L. Winger, and E. Jernigan. Adaptive wiener filtering of noisy images and image sequences. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 3, pages III – 349–52 vol.2, sept. 2003.

[JOP$^+$] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–.

[Jos09] Ebi Jose. *Automatic Feature Detection and Terrain Mapping Using Millimetre Wave Radar for Autonomous Land Navigation*. PhD thesis, School of Electrical & Electronic Engineering, Nanyang Technological University, 2009.

[Kal10] Dinesh Sharma Kalpoe. Vibration measurement of a model wind turbine using high speed photogrammetry. Master's thesis, Delft University of Technology, November 2010.

[KC94] S. Kuttikkad and R. Chellappa. Non-gaussian cfar techniques for target detection in high resolution sar images. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 1, pages 910–914 vol.1, Nov 1994.

[KHW+03] R Krishnan, E Hermann, R Wolff, M Zimmermann, V Seifert, and A. Raabe. Automated fiducial marker detection for patient registration in image-guided neurosurgery. *Computer Aided Surgery*, 8(1):17–23, 2003.

[Kol78] A.N. Kolmogorov. *Stationary sequences in Hilbert space.* John Crerar Library] National Translations Center, 1978.

[Lan96] Dirk Langer. An integrated MMW radar system for outdoor navigation. In *International Conference on Robotics and Automation*, pages 417–422, Minneapolis, Minnesota, April 1996. IEEE.

[Lee80] Jong-Sen Lee. Digital image enhancement and noise filtering by use of local statistics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-2(2):165–168, 1980.

[LKC04] Young Jin Lee, Do-Yoon Kim, and Myung Jin Chung. Feature matching in omnidirectional images with a large sensor motion for map generation of a mobile robot. *Pattern Recognition Letters*, 25(4):413–427, 2004.

[Low04] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[LZ12] Guomin Luo and Daming Zhang. *Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology*, chapter Wavelet Denoising, pages 59 – 80. InTech, 2012.

[MFALRAW07] Marcos Martin-Fernandez, Carlos Alberola-Lopez, Juan Ruiz-Alzola, and Carl-Fredrik Westin. Sequential anisotropic wiener filtering applied to 3d mri data. *Magnetic Resonance Imaging*, 25(2):278 – 292, 2007.

[MKR+10] J. Mullane, S. Keller, A. Rao, M. Adams, A. Yeo, F. Hover, and N. Patrikalakis. X-band radar based SLAM in Singapore's off-shore environment. In *The Eleventh International Conference on Control, Automation, Robotics and Vision (ICARCV 2010)*, Singapore, December 2010.

[MM99] Songrit Maneewongvatana and David M Mount. It's okay to be skinny, if your friends are fat. In *Center for Geometric Computing 4th Annual Workshop on Computational Geometry*, volume 2, pages 1–8, 1999.

[MN03] Niloy J Mitra and An Nguyen. Estimating surface normals in noisy point cloud data. In *Proceedings of the nineteenth annual symposium on Computational geometry*, pages 322–328. ACM, 2003.

[MVAV11] J. Mullane, B.-N. Vo, M. Adams, and B.-T. Vo. *Random Finite Sets for Robot Mapping and SLAM*. Springer Tracts in Advanced Robotics 72. Springer, Berlin Heidelberg, 2011.

[NHY+00] K Nakamura, T Hara, M Yoshida, T Miyahara, and H Ito. Optical frequency domain ranging by a frequency-shifted feedback laser. In *IEEE Journal of Quantum Electronics*, volume 36, pages 305–316, March 2000.

[NS12] T. Nakamura and S. Suzuki. Simplified slam using reflection intensity of laser range sensor with retro-reflective marker. In *2012 Proceedings of SICE Annual Conference (SICE)*, pages 2074–2079, Aug 2012.

[RB03] Søren Riisgaard and Morten Rufus Blas. *SLAM for Dummies (A Tutorial Approach to Simultaneous Localization and Mapping)*. MIT, 2003.

[Roh83] H. Rohling. Radar cfar thresholding in clutter and multiple target situations. *Aerospace and Electronic Systems, IEEE Transactions on*, AES-19(4):608–621, July 1983.

[Ros07] K.H. Rosen. *Discrete Mathematics and Its Applications*. Discrete Mathematics and Its Applications. McGraw-Hill Higher Education, 2007.

[ROSF11] D. Ronzoni, R. Olmi, C. Secchi, and C. Fantuzzi. Agv global localization using indistinguishable artificial landmarks. In *2011 IEEE International Conference on Robotics and Automation*, pages 287–292, May 2011.

[RRKB11] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. Orb: An efficient alternative to sift or surf. In *2011 International Conference on Computer Vision*, pages 2564–2571, Nov 2011.

[SBH+02] S. Scheding, G. Brooker, R. Hennessy, M. Bishop, and A. Maclean. Terrain imaging and perception using millimetre wave radar. In *Interbational Conference on Control, Automation, Robotics and Vision*, Singapore, November 2002.

[SF96] Izhak Saban and Sorin Faibish. Image processing techniques for laser images. In *Electrical and Computer Engineering, 1996. Canadian Conference on*, volume 1, pages 462–465. IEEE, 1996.

[SGB10] Bastian Steder, Giorgio Grisetti, and Wolfram Burgard. Robust place recognition for 3d range data based on point features. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 1400–1405. IEEE, 2010.

[SKC13] J. Seong, J. Kim, and W. Chung. Mobile robot localization using indistinguishable artificial landmarks. In *2013 10th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pages 222–224, Oct 2013.

[Sko08] Merrill I. Skolnik, editor. *Radar Handbook*, chapter Automatic Detection, Tracking, and Sensor Integration. McGraw-Hill, third edition, 2008.

[SS02] L. Shapiro and G. Stockman. *Computer Vision*, chapter 3, Binary Image Analysis, pages 69–73. Prentice Hall, 2002.

[SVV08] D. Schuhmacher, Ba-Tuong Vo, and Ba-Ngu Vo. A consistent metric for performance evaluation of multi-object filters. *Signal Processing, IEEE Transactions on*, 56(8):3447–3457, Aug 2008.

[SWB02] J Schlaier, J Warnat, and A Brawanski. Registration accuracy and practicability of laser-directed surface matching. *Computer Aided Surgery*, 7(5):284–290, 2002.

[SWK07] Ruwen Schnabel, Roland Wahl, and Reinhard Klein. Efficient ransac for point-cloud shape detection. *Computer Graphics Forum*, 26(2):214–226, June 2007.

[TS13] N. Takahiko and S. Suzuki. Kindergarten slam utilizing laser range sensor and retro-reflective markers. In *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, pages 8300–8305, Nov 2013.

[vdWSN+14] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors. scikit-image: image processing in Python. *PeerJ*, 2:e453, 6 2014.

[WCBS+06a] E. Widzyk-Capehart, G. Brooker, S. Scheding, R. Hennessy, A. Maclean, and C. Lobsey. Application of millimetre wave radar sensor to environment mapping in surface mining. In *Control, Automation, Robotics and Vision, 2006. ICARCV '06. 9th International Conference on*, pages 1–6, 2006.

[WCBS+06b] E. Widzyk-Capehart, G.M. Brooker, S.J. Scheding, R.C. Hennessy, A.J. Maclean, and C.R. Lobsey. Application of millimeter wave radar sensor to environment mapping in surface mining. In *ICARCV*, Singapore, Dec 2006.

[WEL07] Matthew R Walter, Ryan M Eustice, and John J Leonard. Exactly sparse extended information filters for feature-based slam. *The International Journal of Robotics Research*, 26(4):335–359, 2007.

[WGM+75] B. Widrow, Jr. Glover, J.R., J.M. McCool, J. Kaunitz, C.S. Williams, R.H. Hearn, J.R. Zeidler, Jr. Eugene Dong, and R.C. Goodlin. Adaptive noise cancelling: Principles and applications. *Proceedings of the IEEE*, 63(12):1692–1716, 1975.

[Wie49] N. Wiener. *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*. Technology press books in science and engineering. Technology Press of the Massachusetts Institute of Technology, 1949.

[Wie64] N. Wiener. *Extrapolation, interpolation, and smoothing of stationary time series*. MIT Press, 1964.

[WJ11] Martin Weinmann and Boris Jutzi. Fully automatic image-based registration of unorganized tls data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, 38:1–6, 2011.

[Wol11] Christian Wolff. Frequency-modulated continuous-wave radar. Web page, 2011.

[WTGBS08] X Wang, C Toth, D Grejner-Brzezinska, and H Sun. Integration of terrestrial laser scanner for ground navigation in gps-challenged environment. In *Proceedings of the XXI Congress of International Society for Photogrammetry and Remote Sensing*, pages 3–11, 2008.

[WWHJ11] Ma Weinmann, Mi Weinmann, Stefan Hinz, and Boris Jutzi. Fast and automatic image-based registration of tls data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(6):S62–S70, 2011.

[Yia02] Peter N. Yianilos. Normalized forms for two common metrics. Technical report, NEC Research Institute, 1991,2002.

[ZC00] Shi Zhong and V. Cherkassky. Image denoising using wavelet thresholding and model selection. In *Image Processing, 2000. Proceedings. 2000 International Conference on*, volume 3, pages 262 –265 vol.3, 2000.

[ZLS05] Fei Zhang, Chengfang Li, and Lina Shi. Detecting and tracking dim moving point target in ir image sequence. *Infrared Physics & Technology*, 46(4):323–328, 2005.

Appendix A

BINR paper

Radar Noise Reduction Based on Binary Integration

Daniel Lühr, *Member, IEEE*, and Martin Adams, *Senior Member, IEEE*

***Abstract*— Short range radars can provide robust information about their surroundings under atmospheric disturbances, such as dust, rain, and snow, conditions under which most other sensing technologies fail. However, this information is corrupted by received power noise, resulting in false alarms, missed detections, and range/bearing uncertainty. The reduction of radar image noise, for human interpretation, as well as the optimal, automatic detection of objects, has been a focus of radar processing algorithms for many years. This paper combines the qualities of the well established binary integration detection method, which manipulates multiple images to improve detection within a static scene, and the noise reduction method of power spectral subtraction. The binary integration method is able to process multiple radar images to provide probability of detection estimates, which accompany each power value received by the radar. The spectral subtraction method then utilizes these probabilities of detection to form an adaptive estimate of the received noise power. This noise power is subtracted from the received power signals, to yield reduced noise radar images. These are compared with state-of-the-art noise reduction methods based on the Wiener filter and wavelet denoising techniques. The presented method exhibits a lower computational complexity than the benchmark approaches and achieves a higher reduction in the noise level. All of the methods are applied to real radar data obtained from a 94-GHz millimetre wave FMCW 2D scanning radar and to synthetic aperture radar data obtained from a publicly available data set.**

***Index Terms*— Binary integration, CFAR, data integration, image denoising, millimeter wave radar, noise reduction, noise subtraction, radar detection, radar imaging, wavelet denoising, Wiener filter, SAR.**

I. Introduction

LANDMARK identification concerns the detection of signals from noisy measurement data. When time is available to obtain multiple images from a static scene at the same location, it is possible to exploit the correlation in the sequence of images to reduce noise, and consequently improve detection. A noise reduction method applied to radar data using these concepts is presented in this work.

Several methods have been developed in the field of image processing to reduce noise in both stationary and dynamic image sequences in applications as diverse as object tracking surveillance, autonomous navigation, motion analysis, and astronomical and medical imaging [1]. A sequence of 2D images is represented by a 3D volume where the third dimension corresponds to the temporal dimension or the sequence index. Many of the methods to process such 3D signals have been developed by generalising well known 2D filtering techniques by extending the support of a filter in the temporal domain. The classical Wiener filter [2], extended to a 3D form, is an example of such an approach. The Wiener filter is a linear time-invariant estimator which adopts a *Minimum mean square error* (MSE) statistical approach. Adaptive noise cancelling, developed as a variation of the original Wiener optimal filtering theory, was presented in [3]. The adaptive noise cancelling application uses a reference signal correlated with the noise to obtain a noise estimate. This estimate is then subtracted from the noisy signal. An application of Wiener filtering to 3D medical imaging data [4] extends the classical Wiener implementation by estimating the filter parameters using a sequence of observations based on the calculation of local statistics (calculated in a small window around the point of interest). In radar applications, a 2D Wiener filter, also using local statistics, has been used to reduce noise in weather radar data [5].

Work by Donoho and Johnstone [6]–[8] introduced the denoising capabilities of the Wavelet transform. The basic method is *Denoising by Thresholding* [9]. It is analogous to frequency domain filtering based on the Fourier transform. The wavelet time-frequency approach however, attempts to reduce noise by preserving a number of coefficients associated with components with high information energy, and discarding the rest. It is assumed that noise (often considered to be additive Gaussian) is spread homogeneously among all signal frequency components. Thus by discarding the coefficients of the components not highly correlated with the signal, a significant amount of noise is eliminated. Coefficients with a magnitude higher than the threshold are considered to hold mostly signal information, and those lower than the threshold are considered to carry mainly noise energy. In radar applications several articles have demonstrated the use of wavelets to reduce noise. Chen [10] proposed a recursive thresholding method for radar image denoising, while Aly demonstrated the use of wavelet packet transforms and higher order statistics to detect and localise RF radar pulses in noisy environments [11]. In general, most noise reduction algorithms (both 'classical' and wavelet-based) assume the noise to be additive Gaussian [1], [9], which is useful for a broad range of applications. However, in radar imaging, the Gaussian noise assumption is often unrealistic. Another critical aspect

Manuscript received June 7, 2014; accepted August 10, 2014. Date of publication August 26, 2014; date of current version November 20, 2014. This work was supported in part by Conicyt Fondecyt Project 1110579, in part by Conicyt-DAAD PCCI-2012009, and in part by AMTC, Universidad de Chile. The associate editor coordinating the review of this paper and approving it for publication was Dr. Lorenzo Lo Monte.

D. Lühr is with the Department of Electrical Engineering, University of Chile, Santiago, Chile (e-mail: dluhr@ieee.org).

M. Adams is with the Department of Electrical Engineering, University of Chile, Santiago, Chile, and also with the Advanced Mining Technology Center, University of Chile, Santiago 1058, Chile (e-mail: martin@ing.uchile.cl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2014.2352295

of wavelet denoising is the appropriate threshold selection. An adaptive threshold method was introduced by Chen [12], which adapts the threshold to the coefficients' statistics, relaxing the Gaussian assumption of most wavelet based methods. Another adaptive method, presented by Jin [13] uses abrupt changes in the signal to adapt weights to calculate local means and variances. This approach is reported to reduce the ripple-like artifacts usually found around edges when using wavelet denoising techniques.

In this work a different approach for radar image denoising is introduced and compared to the classical Wiener filter and to the more recent wavelet based denoising approaches. The proposed method's implementation presents a lower computational complexity than both the 3D Wiener filter and the 3D wavelet approaches. The method further reduces the mean noise value when compared with the other two methods. It uses statistical information provided by the Binary integration detector to identify parts of the received signal corresponding to noise. It uses those parts to obtain an estimate of the noise power spectrum by recursive averaging. This noise estimate is then used for power spectral subtraction [14] (or noise subtraction) to reduce noise. In particular, Binary Integration (BI) combines the output of several single-observation detectors to improve the detection probability, while maintaining the desired, acceptable false alarm rate. The single-observation detector used is a member of the Constant False Alarm Rate (CFAR) family of stochastic detectors, widely used in radar [15].

The following section summarises classical radar detection and demonstrates its application to scanning radar images. Section III presents the three different noise reduction techniques, which will be applied to scanned radar data for comparison purposes. Section IV then explains the implementation details of the methods and analyses the implementation complexity for the three different approaches. Finally, results using millimetre wave (MMW) radar data in an outdoor environment and an open *Synthetic Aperture Radar* SAR data set are also presented in Section V.

II. Radar Detection

Targets of interest in radar data are usually embedded in noise and clutter. Thus, landmark detection is necessary to identify landmark signals from the noisy power measurement data. In this work, statistical information provided by detection methods is used to obtain a noise estimate from a sequence of radar measurements of the same scene. The noise estimate is then used in a noise subtraction method to obtain a reduced noise version of the radar power measurement data.

This section briefly describes the detection methods used, their main equations and parameters and their most important aspects.

Adaptive, stochastic, landmark detection techniques offer principled methods of detection based on a predefined acceptable probability of false alarm and quantifiable probabilities of detection. The Constant False Alarm Rate (CFAR) concept refers to a family of adaptive algorithms widely used in radar to detect target returns against a background of noise, clutter and interference.

In most radar signal processing literature to date, a Cell Averaging (CA) CFAR detector is the preferred method of target detection [16]–[18]. A CA-CFAR processor is used on the experimental data presented in Section V-B.

[19] shows that the detection probability $P_D^{\text{CA-CFAR}}(q)$ of a Rayleigh fluctuating target, embedded in exponential noise or clutter, can be determined from the CA-CFAR parameters

$$P_D^{\text{CA-CFAR}}(q) = \left[1 + \frac{\tau^{\text{CA-CFAR}}}{W_f}\left(\frac{1}{1+\widehat{\eta}^{\text{SNP}}(q)}\right)\right]^{-W_f} \quad (1)$$

where W_f is the size of the CFAR window and $\tau^{\text{CA-CFAR}}$ is defined as

$$\tau^{\text{CA-CFAR}} = W_f\left(\left(P_{fa}^{\text{CA-CFAR}}\right)^{\frac{-1}{W_f}} - 1\right) \quad (2)$$

and $\widehat{\eta}^{\text{SNP}}(q)$ is the estimated received SNP calculated as

$$\widehat{\eta}^{\text{SNP}}(q) = \frac{S_{\text{lin}}^{\text{radar}}(q)}{T(S_{\text{lin}}^{\text{radar}}(q))} \quad (3)$$

where $S_{\text{lin}}^{\text{radar}}(q)$ is the linearised received power from the radar in the q^{th} (bearing-range) bin, and $T(S_{\text{lin}}^{\text{radar}}(q))$ is the CFAR test statistic, which in the case of CA-CFAR, corresponds to the sample mean of the neighbouring cell's power values in the CFAR window. The adaptive threshold is then defined as

$$S_{\text{lin}}^{\text{CA-CFAR}}(q) = \tau^{\text{CA-CFAR}} \cdot T(S_{\text{lin}}^{\text{radar}}(q)). \quad (4)$$

Several other CFAR methods have been developed and current research focuses in CFAR methods with adaptive parameters [20]. In particular, the Ordered Statistic (OS) CFAR has been reported to perform well for large targets (with respect to the spatial resolution) and in SAR images, due to their noise and clutter being usually modelled by Weibull or K distributions, and the higher effect of multiplicative speckle noise present in such images. The OS-CFAR method is used on the SAR data set presented in Section V-C.

In the OS-CFAR method the test statistic $T(S_{\text{lin}}^{\text{radar}}(q))$ is obtained by choosing the k^{th} value from the ordered set of power values in the CFAR window

$$S_{\text{lin}}^{(1)} \le S_{\text{lin}}^{(2)} \le \ldots \le S_{\text{lin}}^{(k)} \le \ldots \le S_{\text{lin}}^{(W_f-1)} \le S_{\text{lin}}^{(W_f)} \quad (5)$$

a value of $k = \frac{3W_f}{4}$ has been suggested in [21] to represent a good estimate for typical radar applications. The parameter $\tau^{\text{OS-CFAR}}$ needs to be calculated numerically from

$$P_{fa}^{\text{OS-CFAR}} = \prod_{i=0}^{k-1} \frac{W_f - i}{W_f - i + \tau^{\text{OS-CFAR}}} \quad (6)$$

where $P_{fa}^{\text{OS-CFAR}}$ is the chosen acceptable OS-CFAR probability of false alarm. The probability of detection, $P_D^{\text{OS-CFAR}}(q)$, is obtained from

$$P_D^{\text{OS-CFAR}}(q) = \prod_{i=0}^{k-1} \frac{W_f - i}{W_f - i + \frac{\tau^{\text{OS-CFAR}}}{1+\widehat{\eta}^{\text{SNP}}(q)}} \quad (7)$$

Unfortunately, the noise and target distribution assumptions in CFAR are often violated, yielding higher false alarm

and missed detection rates than those theoretically derived. To reduce this problem, and if time is available to acquire multiple scans of a fixed environment, it makes sense to exploit the high target correlation between scans to further reduce the uncertainty in the existence of objects and reduce the noise. Techniques which implement this concept are generally referred to as integration methods. A simple but effective method widely used in the radar community is the Binary Integration (BI) Method [19].

When integrating L scans, the probability of detection yielded by this method is

$$P_D^{\mathrm{BI}} = \sum_{j=M^{\mathrm{BI}}}^{L} \frac{L!}{j!(L-j)!} \left(P_D^{\text{CA-CFAR}}\right)^j \left(1 - P_D^{\text{CA-CFAR}}\right)^{L-j} \tag{8}$$

where $P_D^{\text{CA-CFAR}}$ is the probability of detection in a *single scan*, and $M^{\mathrm{BI}} < L$ is the optimal BI parameter for a given L. Likewise, if $P_{fa}^{\text{CA-CFAR}}$ is the probability of false alarm in a single scan, then the probability of false alarm for the binary integration method is

$$P_{fa}^{\mathrm{BI}} = \sum_{j=M^{\mathrm{BI}}}^{L} \frac{L!}{j!(L-j)!} \left(P_{fa}^{\text{CA-CFAR}}\right)^j \left(1 - P_{fa}^{\text{CA-CFAR}}\right)^{L-j} \tag{9}$$

BI offers a robust technique to identify which parts of signal are noise. The detection probability obtained by means of the BI method will be used as a target existence probability estimate in the radar noise reduction method presented in section III-C.

III. Noise Reduction Methods

In this section three different noise reduction approaches are described. The first state of the art method is based on a version of the linear Wiener filter which uses estimates of the local means and variances in order to estimate the noise characteristics. The second, more recent, non-linear method, is a wavelet denoising approach. The final method, proposed in this article, is based on spectral noise subtraction. This method, is capable of preserving non-linear features.[1]

A. Wiener Filter

This method of noise reduction corresponds to the application of a discrete-time minimum-mean-square-error filter. Such a filter is known as a Wiener-Kolmogorov filter or Wiener filter for short [22], [23].

Wiener filters assume additive noise and that the signals are stationary, linear stochastic processes. Because radar images, as well as natural images, consist of smooth areas, textures and edges, they are not *globally* stationary, but can be treated as *locally* stationary. This led to the derivation of the Lee filter [24], which has been extensively used in video denoising, where it has proved to be successful in terms of noise removal and preserving some important image features (e.g. edges) [13]. The Lee filter assumes that all samples within a local window are from the same structure (local stationarity). This assumption is invalid when sharp edges are encountered within the window, therefore the mean is blurred and the variance increases near the edges, which results in a degraded image near those regions.

B. Wavelet Denoising

Some of the limitations of the Wiener filter, particularly its inability of preserving non-linear features in the data, can be overcome by using non-linear filters. However, in general, finding the parameters for a non-linear filter is a complex task. Since the introduction of the wavelet denoising methods by Donoho and Johnstone [6]–[8] a powerful, yet simple to implement non-linear filter for noise reduction has become available and is widely used in practical implementations.

The particular thresholding function used in this work is the universal threshold proposed by Donoho and Johnstone [25], [26] with the soft-thresholding method proposed in [6].

This method generates an estimated signal with a smaller amplitude than the original one, but it retains the regularity[2] of the signal.

C. Spectral Noise Subtraction

Noise subtraction methods were originally devised for noise reduction in noisy speech signals [27]. In the case of radar data, the binary integrator's probability of detection can be used to identify sections which have low probability of having any target information and therefore they can be used to estimate the noise magnitude.

The noise power estimate $\widehat{\Sigma}_n(l)$, when the l^{th} observation has been received, can be calculated as in Equation (10), which corresponds to a recursive smoother using a fixed parameter α_d, and the binary integration probability of detection from Equation (8).

$$\begin{aligned}\widehat{\Sigma}_n(l) = &\left(\alpha_d \widehat{\Sigma}_n(l-1) + (1-\alpha_d) S_{\mathrm{lin}}^{\mathrm{radar}}(l)\right)\left(1 - P_D^{\mathrm{BI}}(l)\right) \\ &+ \widehat{\Sigma}_n(l-1) \times P_D^{\mathrm{BI}}(l)\end{aligned} \tag{10}$$

where l corresponds to the observation number. The first term on the right of Equation (10) represents the smoothed (averaged) noise power, weighted by $\left(1 - P_D^{\mathrm{BI}}(l)\right)$ during *target absence* sections of the signal, while the second term shows that the previous estimate is preserved and not updated if there is a high probability of *target presence* $\left(P_D^{\mathrm{BI}}(l) \to 1\right)$. Introducing

$$\tilde{\alpha}_d(l) = \alpha_d + (1-\alpha_d) P_D^{\mathrm{BI}}(l) \tag{11}$$

Equation (10) can be rewritten as

$$\widehat{\Sigma}_n(l) = \tilde{\alpha}_d(l) \widehat{\Sigma}_n(l-1) + \left[1 - \tilde{\alpha}_d(l)\right] S_{\mathrm{lin}}^{\mathrm{radar}}(l) \tag{12}$$

In Equation (11), $\tilde{\alpha}_d(l)$ is a time varying smoothing parameter. Hence the noise spectrum can be estimated using past

[1] non-linear features refer, in general, to sharp edges or discontinuities within a signal. E.g., in radar data, the sudden change in received power caused by the presence of a target.

[2] regularity corresponds to areas in which the signal is continuous, while discontinuities correspond to irregular points.

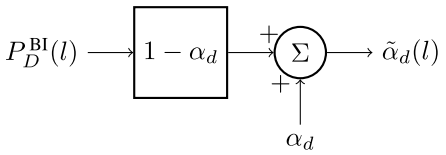

Fig. 1. Block diagram summarizing the time varying parameter $\tilde{\alpha}_d$.

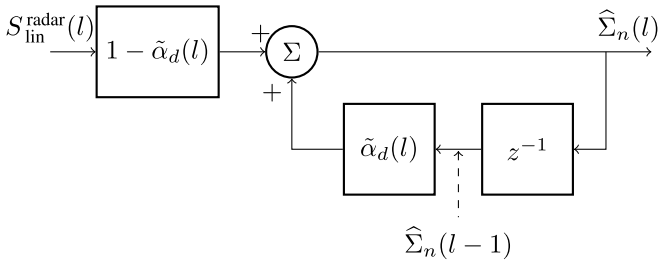

Fig. 2. Binary integration noise reduction block diagram.

spectral received power values, together with a smoothing parameter which itself varies according to the BI probability of detection $P_D^{\mathrm{BI}}(l)$. The adaptive noise estimator is summarised in the block diagrams of Figures 1 and 2.

The noise power estimate $\widehat{\Sigma}_n(l)$ can be used to obtain a *Binary Integration Noise Reduction* (BINR) power estimate $\hat{S}_{\mathrm{lin}}^{\mathrm{BINR}}(l)$ using the method of power spectral subtraction [14]. In the basic spectral subtraction algorithm, the average noise power, $\widehat{\Sigma}_n(l)$ is subtracted from the noisy range spectrum. In [28], a method which further reduces background noise for SNPs between -5 and 20 dB was devised based on subtracting an over-estimate of the noise power and preventing the resultant spectral components from reaching below a preset minimum level, termed the "spectral floor". This method leads to a reduced noise power estimate $\hat{S}_{\mathrm{lin}}^{\mathrm{BINR}}(l)$ given by

$$\widehat{S}_{\mathrm{lin}}^{\mathrm{BINR}}(l) = \begin{cases} S_{\mathrm{lin}}^{\mathrm{radar}}(l) - c\,\widehat{\Sigma}_n(l) & \text{if } S_{\mathrm{lin}}^{\mathrm{radar}}(l) > c\,\widehat{\Sigma}_n(l) \\ d \times \widehat{\Sigma}_n(l) & \text{otherwise} \end{cases} \tag{13}$$

where c is an over-subtraction factor ($c \geq 1$) and d is the spectral floor parameter ($0 < d < 1$). A value of c which is larger than 1 represents the fact that it is necessary to subtract more than the expected value of the noise (which is a smoothed estimate) to make sure that most of the noise peaks are removed. The spectral floor parameter d, when greater than zero, ensures that the remnants of the noise peaks are masked by neighbouring spectral components of comparable magnitude. This results in a reduction of broadband noise, when compared with the original power spectral subtraction method. [28] further demonstrated the advantages of an adaptive over-subtraction factor c, which varies between frames of recorded spectra, as a function of the estimated signal.

IV. Implementation and Computational Complexity

This section discusses the implementation and computational complexity, as a function of the data size, of all three noise reduction methods. All the algorithms were implemented using the SciPy signal processing toolbox [29]. The same received radar power signal S is used as the input for all the methods. The signal includes additive noise. It is a 3D array formed by stacking L radar scans. Each radar scan $S(l)$ corresponds to a B-scope of size $N_p = N_r \times N_b$ cells (pixels), where N_r represents the number of range bins (rows) while N_b the number of bearing bins (columns). The first dimension in the 3D array represents range (r), the second represents bearing (b) and the third represents the scan number (l). The sequence of L power values for a particular range and bearing is denoted as $S_{r,b}$ while an individual voxel is referred to as $S_{r,b,l}$. The total number of voxels in S is $N_v = N_b \times N_r \times L$. The complexity analysis will be carried out in terms of N_p and L ($N_v = N_p \times L$), to separate the effects of the radar image size from those related to the number of observations.

A. Wiener Filter

The Wiener filter noise reduction technique used in this work requires a support volume, in order to calculate the local means and variances. For a given voxel, the cells contained in the support volume around it will be used to calculate the local statistics. Let $(2K_r+1)$, $(2K_b+1)$, $(2K_l+1)$ be the dimensions of this volume in the range (rows), bearing (columns) and observation (depth) dimensions. K_r and K_b are related to the expected range and bearing size of targets. If these parameters are set too high the local mean and variance of a target voxel will incorporate not only target values but also undesired noise values. On the other hand, if they are to small, the local mean and variance estimates would be degraded. $2K_l + 1 = L$ has been chosen for the window to incorporate information from all observations at each value in the temporal (l) dimension. This algorithm exhibits a complexity of the order $\mathcal{O}(N_p \times L^2)$.

B. Wavelet Denoising

The 3D wavelet denoising method has been separated into two sequential wavelet denoising problems. The first carries out scan by scan 2D spatial wavelet denoising, while the second executes 1D time domain wavelet denoising for each range-bearing cell sequence. General wavelet methods have a complexity of order $\mathcal{O}(N \log N)$, where N corresponds to the data size (total number of pixels for 2D, or signal length for the 1D case). Therefore, the first part is of the order $\mathcal{O}(L \times N_p \log N_p)$ while the second part presents a complexity of the order $\mathcal{O}(N_p \times L \log L)$. Hence the total complexity of the wavelet method is of the order $\mathcal{O}(L \times N_p \log N_p + N_p \times L \log L)$ which reduces to $\mathcal{O}\big((N_p \times L)\log(N_p \times L)\big)$.

C. Binary Integration Noise Reduction

The optimum parameters required for Binary Integration, $M_{\mathrm{opt}}^{\mathrm{BI}}$ and $P_{fa}^{\mathrm{CA\text{-}CFAR}}$, can be calculated off-line and they are not included in the algorithm, but given as inputs. Also, the BI probability of detection, which is dependent on $M_{\mathrm{opt}}^{\mathrm{BI}}$ and L, can also be defined offline, and the corresponding polynomial in $P_{D_{r,b,l}}^{\mathrm{CA\text{-}CFAR}}$ is therefore, pre-calculated. The combination of the CA-CFAR, BI and noise subtraction parts of the algorithm yields a complexity of the order $\mathcal{O}(N_p \times L)$.

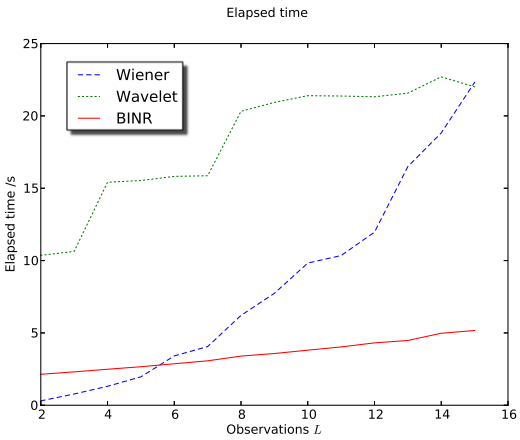

Fig. 3. Elapsed computational time measurement.

Fig. 4. Park environment where radar data was captured (obtained from Google Earth).

In summary, with respect to observations L, the Wiener filter algorithm has the highest complexity (quadratic), followed by the linearithmic[3] complexity of the Wavelet approach, while the binary integration noise reduction's linear complexity makes it the least complex. On the other hand, the Wiener filter and the binary integration noise reduction are the least complex (linear time) with respect to the data size N_b.

V. Results

All of the noise reduction methods have been tested with real data obtained in a local park environment and with an open SAR image data set. The results show reduced noise radar power in PPI representation, as well as, the average noise level versus observation number. In order to demonstrate the usefulness of each reduced noise data set, the CFAR detection method is finally applied to each set. CA-CFAR is used in the park data set due to its compliance to the detector requirements as described in Section II, while OS-CFAR is used in the SAR data set, which has been proven to be more effective with this kind of data [34]. The reduced noise CFAR output is then compared to the CFAR detector's result on the original noisy data. The computational times of the algorithms are also compared.

A. Computational Time

The computational time used by the different algorithms, plotted against observation number L, is shown in Figure 3. The results are consistent with the analysis presented in section IV. The Wiener filter's complexity grows approximately quadratically with L, while the wavelet exhibits linearithmic complexity. A linear time complexity is achieved by the BINR method.

B. Experimental Data

An experimental radar data set, captured[4] in a public park in Santiago, is used to test the noise reduction schemes.

[3]A *linearithmic* function is of the form $n \log n$. An algorithm with a time complexity of the order $\mathcal{O}(n \log n)$ is said to run in *linearithmic* time.

[4]using an Acumine 94 GHz, scanning radar [31].

TABLE I
Optimal M^{BI} Parameter for Different Number of Observations L

L	2	5	10	15	20
M^{BI}	1	2	4	5	7

1) Noise Reduction: Noise values in real radar data do not conform to perfect Gaussian or exponential distributions, as assumed by the noise reduction methods, which impairs their performance. An analysis of the noise reduction methods considering the park environment now follows. Although the methods were applied to the B-Scope radar data (range vs. bearing), the results are shown in *plan position indicator* (PPI) form for clearer visualisation. The test environment is shown in Figure 4. The area corresponds to a main paved track approximately 65 m wide. On the sides of the track there are lamp posts and some trees. There are also fences and concrete walls. The radar was located on the track.

The CA-CFAR window size was 9 bins in the bearing direction and 7 bins in the range direction. The guard cells window size was 5 in the bearing dimension and 3 bins in the range direction. These parameters were found suitable, in preliminary experiments, for detecting the lamp posts and trees surrounding the radar, by considering the power spread these features produce in the acquired data.

The BI false alarm rate used was 1×10^{-6}. The optimal M^{BI} has been previously obtained for different L values. Some of the values are listed in Table I. The results presented correspond to $L = 20$ observations. For the noise subtraction algorithm, the chosen parameters were $\alpha_d = 0.9$, $c = 50.0$ and $d = 0.1$. A high α_d ensures that the previous value of the noise estimate has more weight than the new observation which is desirable given the high noise levels present in radar data. Parameter c, controlling over-subtraction, was selected by testing different values between 10 and 100. Similarly, the spectral floor parameter d was tested for different values between 0.05 to 0.5, with the chosen value yielding good results in the reduction of the broadband noise.

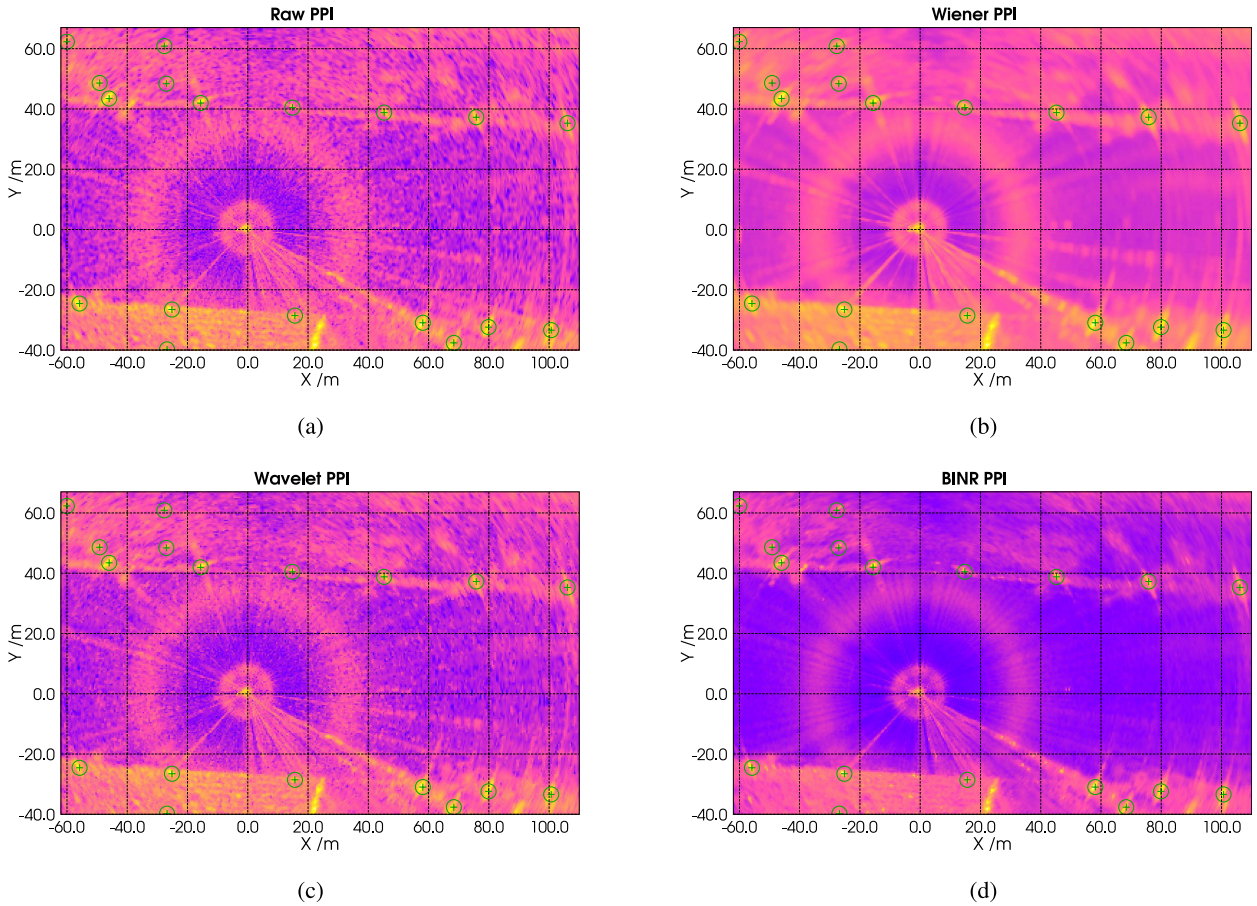

Fig. 5. Raw power and reduced noise power PPI plots of the park area. (a) PPI showing noisy input data from the park environment. (b) PPI showing Wiener filtered data from the park environment. (c) PPI showing Wavelet denoised data from the park environment. (d) PPI showing BINR data from the park environment.

The Wiener support region was 3 bins in the range and bearing direction, which is the expected power spread of the targets of interest.

In the case of the spatial (2D) wavelet denoising, the Daubechies 3 wavelet function was used, while the Haar wavelet was selected for the 1D (temporal) dimension.

The noisy raw radar input data from the park is presented in Figure 5a. The ground truth location of lamp posts and trees are marked with green circles and a cross in their centre.

Wiener filtering (Figure 5b) exhibits a smoother noise background but the main objects identified in the scene are blurred by the filter, thus losing localisation detail. Wavelet denoising (Figure 5c), is able to preserve the location and edges of targets. It does, however, produce several negative values in noise only sections, which are truncated to a small value to allow visualisation. Nevertheless, the average noise level is reduced. Finally, the BINR method in Figure 5d shows its ability to retain details as well as to reduce the noise level. It can, however, be observed that some of the maximum power peaks have been reduced in magnitude (e.g. tree at $(-49.1$ m, 48.6 m); with a raw power value of 86.02 dB and a BINR power value of 82.73 dB). This is due to the fact that at some observation l those particular targets were not detected, therefore their power values were considered noise and thus subtracted from them. Note that wavelet denoising (Figure 5c) reduces some noise-only areas to very low values. However, the noise background is not homogeneous, therefore, the sharp edges between noise areas near the average noise level and those greatly reduced by the wavelet method can yield several false detections as will be shown, after applying the CA-CFAR detector.

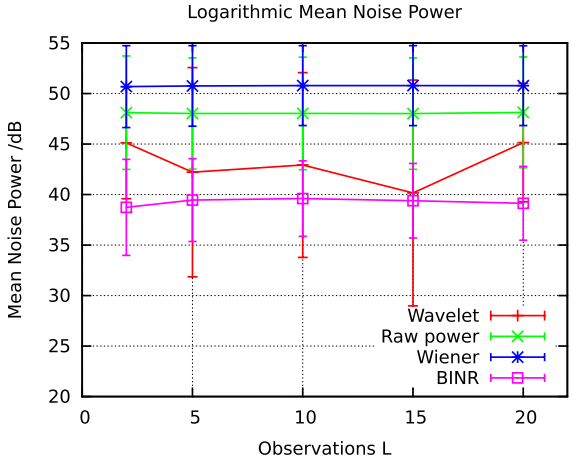

Fig. 6. Noise mean values and variances from a noise only area.

Figure 6 shows mean noise power values from each method, in an area which is known to contain no targets. The BINR

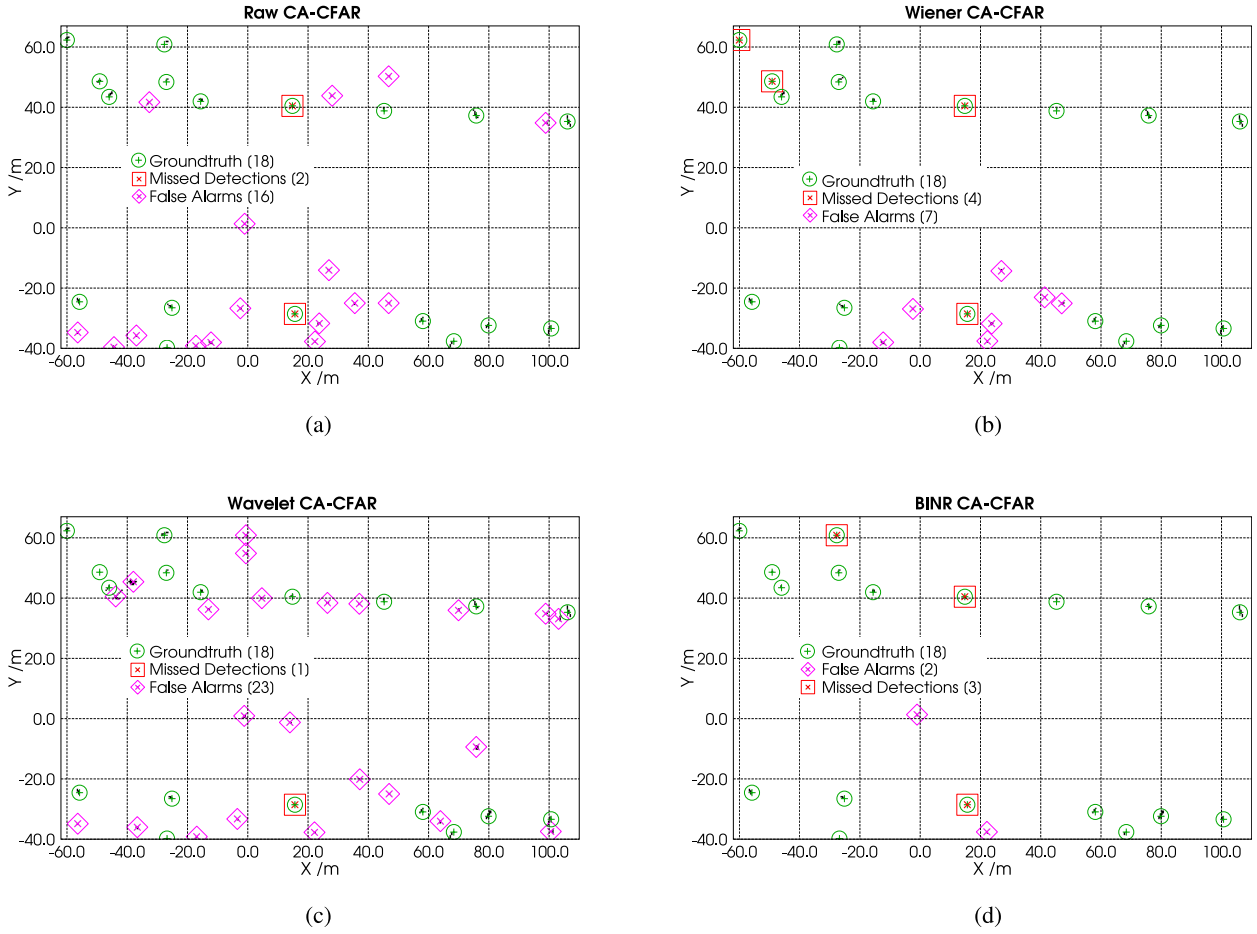

Fig. 7. CA-CFAR detector applied to raw and reduced noise power in the park area. (a) CA-CFAR PPI showing noisy input data from park environment. (b) CA-CFAR PPI showing Wiener filtered data from park environment. (c) CA-CFAR PPI showing Wavelet denoised data from park environment. (d) CA-CFAR PPI showing BINR data from the park environment.

method exhibits the the lowest mean noise power for all L values. Its variance, on the other hand, is higher than that of the Wiener for small L values, but as more observations are included, BINR achieves also the lowest noise variance. The Wiener filter mean noise power stays above that of the raw data, but keeps the variance at a low value. The Wavelet denoising method is able to reduce the mean noise power lower than the raw data but is not able to reach the value of the BINR method. The Wavelet's variance remains high and oscillates due to the ripple effect mentioned before. The noise assumptions which form the basis of all three noise reduction techniques are violated in practice. In particular, the Wiener filter is not optimal for non-Gaussian noise distributions. In the case of the wavelet method, the universal threshold is not able to correctly estimate a noise threshold to separate the noise and information based wavelet coefficients. Furthermore, noise information is no longer spread homogeneously across all wavelet coefficients. Likewise, the CFAR method used in the BINR yields a higher false alarm rate than expected, since exponential noise is assumed, as will be shown in the next section.

2) Target Detection: The CA-CFAR detector is applied to the reduced noise data, in order to demonstrate the usefulness of each noise reduction method. In this case a more relaxed CA-CFAR probability of false alarm is applied ($P_{fa} = 1 \times 10^{-3}$), which reduces $\tau^{\text{CA-CFAR}}$ and increments the $P_D^{\text{CA-CFAR}}$ (see Equations (1) and (2)). This is possible since the reduced noise data is expected to yield a lower false alarm rate, and it helps to overcome the lower target power values obtained in all three methods due to smoothing, thresholding or noise subtraction.

Figure 7a shows the result of the detector applied to the raw noisy input data. The detector itself is able to reduce false alarms to some extent in the raw data. Again, the green circles in the figure denote the ground truth location of some relevant targets (trees and lamp posts at the side of the track) obtained using multiple scans from a laser scanner. CA-CFAR detections are marked with black dots, when these dots are in the vicinity of a ground truth marker, they are considered correct detections. Red square markers are used to show missed detections, while magenta coloured diamonds are used to mark false alarms. The Wiener filter (Figure 7b), interestingly, shows a remarkable reduction of background noise, but its blurring effects present some problems. In particular, small and medium size targets with low SNP are blurred in such a way that their size is reduced and even completely removed from the CA-CFAR detector output. On the other hand, high SNP targets become increased in size (more pixels surrounding

TABLE II
A POSTERIORI DETECTION AND FALSE ALARM RATES

Method	T_D	T_G	T_{fa}	T_N	R_D	R_{fa} /10^{-3}
Raw	16	18	23	81919	0.89	0.2807
Wiener	14	18	14	81943	0.78	0.1708
Wavelet	17	18	88	81712	0.94	1.0758
BINR	15	18	2	81977	0.83	0.0244

TABLE III
OSPA METRIC APPLIED TO EACH METHOD

L	Raw power	Wiener	Wavelet	BINR
2	1.6747	1.6225	1.6565	1.5401
5	1.7055	1.6174	1.8435	1.5230
10	1.6770	1.6377	1.7328	1.4486
15	1.7013	1.6230	1.8673	1.4365
20	1.6796	1.6025	1.7878	1.5114

the landmark are marked as detections), for instance, the tree at coordinates (−27.0 m, 48.4 m). In general, edge details are lost. Figure 7c presents the results of applying the detector to the wavelet denoised data. The method, is able to keep feature details but it shows more false alarms than expected, especially in broad "noise-only" areas. This is due to the noise background not being homogeneous, which leads to the ripple effect of the chosen wavelet function, and the universal threshold not being able to correctly discriminate between information and noise wavelet coefficients. The BINR method (Figure 7d), is able to preserve localisation details, e.g. targets at (−25.1 m, −26.6 m), (68.4 m, −37.6 m) and (100.7 m, −33.4 m), as well as reduce the number of false alarms when compared to the detector applied to raw data and the other noise reduction methods.

A posteriori detection and false alarm rates can be derived from the results. Detection rate, R_D, is obtained by dividing the number of correctly detected ground truth targets, T_D, and the total number of ground truth targets, T_G, i.e. $R_D = T_d/T_G$. The False alarm rate, R_{fa}, on the other hand, is calculated by dividing the number of pixels corresponding to false alarms, T_{fa}, and the sum of pixels correctly identified as noise, T_N, plus the false alarm pixels, i.e. $R_{fa} = T_{fa}/(T_{fa} + T_N)$. Table II summarises the detection and false alarm rates calculation. Note that the area, where the experiment was carried out, exhibits a lower false alarm rate than expected (0.281×10^{-3}), with a detection rate of 89%. The Wiener filter yields a lower false alarm rate (0.171×10^{-3}) but in the same order of magnitude. The detection rate is also lower (78%) than in the raw image. Wavelet denoising's detection rate is higher (94%), but its false alarm rate is higher by an order of magnitude (1.076×10^{-3}). BINR is able to reduce the false alarm rate by an order of magnitude (0.024×10^{-3}), with a detection rate (83%) higher than the Wiener filter, but lower than the raw data. From this result, it can be seen that using BINR it is safe to increase the $P_{fa}^{\text{CA-CFAR}}$ value in order to get a higher detection rate while still achieving a low false alarm rate.

3) Quantifying Detection Performance: While there are several metrics to quantify the error between detected and true targets, a metric is now applied (which jointly considers errors in target location, and number estimates), based on a p^{th} order Wasserstein construction [30]. This metric has been recently introduced and has been shown to produce more consistent results than others (Hausdorff, OMAT) in multi-object systems. It is formulated in terms of the ground truth set of targets $\widehat{\mathcal{M}}$ and the detected set $\mathcal{M}$. If $|\mathcal{M}| > |\widehat{\mathcal{M}}|$, it is given by,

$$\bar{d}^{(\upsilon)}(\widehat{\mathcal{M}}, \mathcal{M}) = \left(\frac{1}{|\mathcal{M}|}\left(\min_{\pi \in \Pi_k} \sum_{i=1}^{|\widehat{\mathcal{M}}|} d^{(\upsilon)}\left(\widehat{m}^i, m^{\pi(i)}\right)^p + \upsilon^p\left(|\mathcal{M}| - |\widehat{\mathcal{M}}|\right)\right)\right)^{1/p} \tag{14}$$

where π corresponds to the permutation in Π_k (the set of permutations on $\{1, 2, \ldots, k\}$) which minimises the sum of the distances between the elements from $\widehat{\mathcal{M}}$ and $\mathcal{M}$, while

$$d^{(\upsilon)}(\widehat{m}^i, m^{\pi(i)}) = \min(\upsilon, ||\widehat{m}^i - m^{\pi(i)}||) \tag{15}$$

is the minimum of the cut-off parameter, υ, and the Euclidean distance between the ground truth target location, $\widehat{m}^i$ and the detected target location $m^{\pi(i)}$. If $|\mathcal{M}| < |\widehat{\mathcal{M}}|$ the metric is obtained through $\bar{d}^{(\upsilon)}(\mathcal{M}, \widehat{\mathcal{M}})$. To find the optimal $\pi \in \Pi_k$ a special case of integer linear programming is required, which can be solved efficiently using the Hungarian method for optimal point assignment.

This metric was applied to the CA-CFAR output of the reduced noise images to quantify the performance of each method. The results for different L values are shown in Table III. Because of the range and angular resolution of the radar used in this experiment, for each ground truth landmark, more than one point is marked as a detection. This increases the error measured by the metric because extra points are considered as false alarms. Thus, the metric output starts to converge to the cut-off parameter, as the second term to the right of Equation (14) dominates. A method to reduce this effect is to combine multiple detections in close proximity using a clustering method, e.g. *connected component labelling* [32]. A different approach, consisting of duplicating the elements of the ground truth set, was used in this implementation, which effectively reduces the effect described above, without altering the results. For all L values, the BINR methods yields a smaller value of the metric meaning that the produced map is closer to the ground truth than that generated by the raw data and the other methods. The Wiener filter is also capable of producing a map with a lower error than the raw data. On the other hand, the Wavelet denoising shows a poor performance due to the false alarms in the background noise caused by the ripple effect within the Wavelet transform.

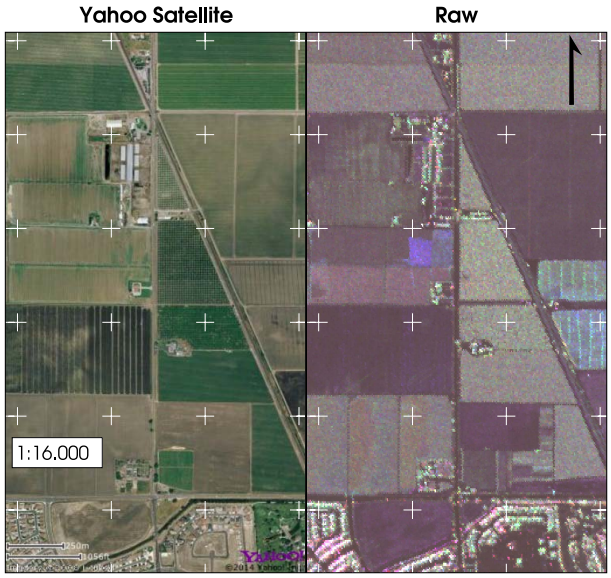

Fig. 8. Yahoo Satellite view (left) and Raw SAR image of the area (right).

C. UAVSAR Data Set

Detection and noise reduction methods in radar are not only used in classical A-Scopes, B-Scopes and PPIs, they can also be used in other forms of radar data such as SAR images [20], [33].

SAR images, being constructed in a fundamentally different way than the classical radar images, are also affected by noise in a different way. In SAR images, noise and clutter are usually modelled by a Weibull or K distribution. Also, the effect of multiplicative speckle noise in SAR images is higher than in other forms of radar data. Under these conditions, the Ordered Statistics (OS) CFAR detection method has proven to be effective when applied to SAR images [34].

In this section the results of using the BINR method on a set of SAR images obtained from the NASA Jet Propulsion Lab (JPL)'s Uninhabited Aerial Vehicle SAR (UAVSAR) mission[5] are presented.

The images correspond to a location near Sacramento, CA, which covers an area of crop fields with isolated buildings in the north-most part (top) of the image and a suburban area with high density housing in the south-most part (bottom). Figure 8 (left) shows a Yahoo Satellite image of the area, with its corresponding SAR image (right). The area is 1.6 km in the horizontal (east-west) direction and 2.88 km in the vertical (north-south) direction. These SAR images represent backscatterred radar power, polarised in the HH, HV and VV components. The magnitude of each component is encoded in the image's red, green and blue channels, respectively.

BINR based on the OS-CFAR detector has been used to first reduce the noise in a series of multiple ($L = 6, M^{\mathrm{BI}} = 3$) observations of the same area. Then the OS-CFAR detector is applied to the reduced noise data to detect buildings. The OS-CFAR window size was 7 bins in the x and y coordinates, while the guard cells window size was 3 bins in both directions. The threshold's constant parameter chosen was $\tau^{\text{OS-CFAR}} = 4.16707$, this value is obtained by solving Eq. (6) using the OS-CFAR window size and the desired value for $P_{fa}^{\text{OS-CFAR}}$. Finally, the noise subtraction parameters used were $\alpha_d = 0.9$, $c = 50.0$ and $d = 0.1$.

[5]UAVSAR data courtesy NASA/JPL-Caltech. http://uavsar.jpl.nasa.gov/

TABLE IV
MEAN NOISE POWER AND VARIANCE IN NOISE ONLY AREA IN dB

	HH		HV		VV	
Method	Mean	Var.	Mean	Var.	Mean	Var.
Raw	24.96	160.85	29.14	350.91	28.62	169.56
Wiener	27.58	159.20	36.59	437.78	32.74	180.68
Wavelet	-13.64	1434.88	-9.62	916.63	-9.75	1042.87
BINR	14.38	166.84	20.30	383.10	19.26	204.06

The parameters for Wavelet denoising and Wiener filtering were the same as those used in the experimental data set presented in Section V-B.

All three polarisation components have been processed. Buildings, in general, reflect radar waves similarly in all polarisations while vegetation and other terrain considered to be clutter in this case, usually exhibit different back-scatter intensity at the different polarisations. The raw power (left) and the output of the OS-CFAR detector are shown in Figure 9a. In the OS-CFAR image, the red, green and blue pixels corresponds to detections in the HH, HV and VV polarisations, respectively. Cyan, magenta and yellow pixels represent detections in the respective combinations of two polarisations, while white pixels represent detections in all three polarisations. Buildings appear in the OS-CFAR image with white pixels (detections in all polarisations), while parks and crop fields present detections in single polarisations or no detection at all.

The Wiener filtering results are presented in Figure 9b. The reduced noise image appears blurred, as expected from the Wiener filter, and the OS-CFAR detector is unable to detect buildings from the crop fields.

The wavelet denoising output is shown in Figure 9c. The power image shows darker colours meaning that the average noise power has been reduced, but several areas present a high variance, particularly in the crop fields. The OS-CFAR output confirms this, and the detector is unable to detect building structures.

Figure 9d corresponds to BINR output. It can be observed that the areas corresponding to crop fields appear smoothed when compared to the raw image. It can also be observed that the number of detections in single polarisations, mostly located in areas corresponding to crop fields, is reduced in the BINR image. On the other hand, most pixels corresponding to building like structures are preserved.

In this data set it is not possible to apply the OSPA metric as the real ground truth is unavailable. An analysis on the noise statistics in an area with no targets, as was carried out in the park data set, quantifies the performance of the noise reduction methods. Table IV shows the mean noise power and variance per polarisation channel. It can be observed that the

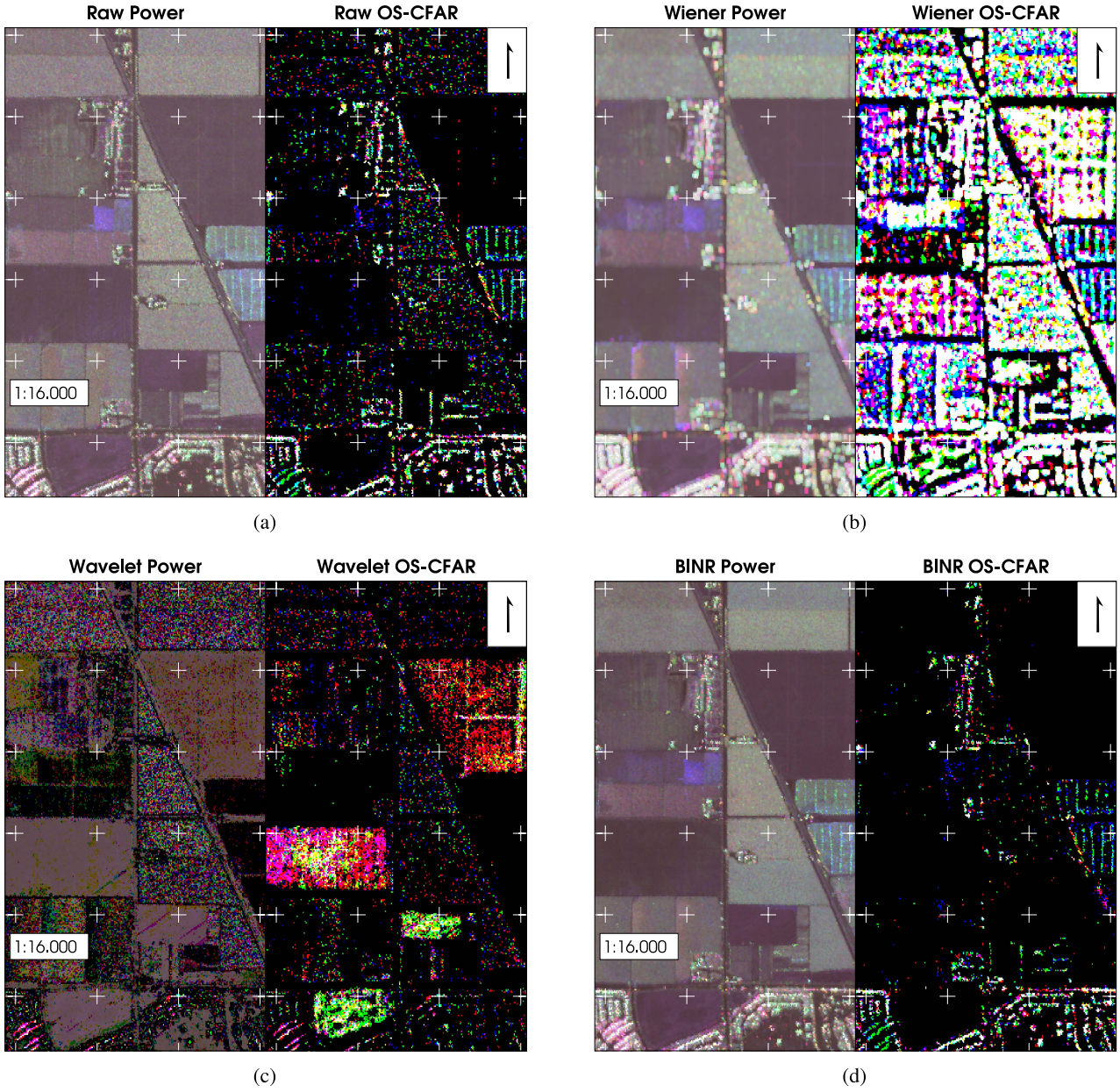

Fig. 9. Raw and reduced noise power SAR images and OS-CFAR applied to them. (a) Raw power SAR image (left) and OS-CFAR applied to it (right). (b) Wiener reduced noise SAR image (left) and OS-CFAR applied to it (right). (c) Wavelet reduced noise SAR image (left) and OS-CFAR applied to it (right). (d) BINR SAR image and OS-CFAR applied to it.

Wavelet presents the lowest mean noise power, but at the same time it yields a very high variance. The Wiener and BINR methods keep a variance similar to that of the original data, but BINR is able to reduce the mean noise level considerably, while the Wiener filter noise level stays at a similar or higher value than the raw data.

VI. Conclusion

In this work, a noise reduction method based on the binary integrator detector and spectral noise subtraction has been introduced. Its performance has been compared with two other general, widely used, noise reduction methods, Wiener filtering and wavelet denoising.

As detailed in Section V, the BINR method exhibits good noise reduction capabilities, which result in a higher signal-to-noise power ratio than the other two techniques. Importantly, BINR has the ability to preserve most of the signal's spatial details. Also, the BINR algorithm's computational time is lower than the other two methods. The reduced time complexity shown by the BINR method and its denoising capabilities make it an appealing noise reduction alternative to be used as a pre-processing step for radar data.

The encouraging results obtained open the door for several improvements for radar data noise reduction. One track for improvement lies in enhancing the standard Wiener or Wavelet methods by making them adaptable to different noise distributions. This approach has been adopted for some

wavelet denoising applications in which different threshold functions are derived for some non-Gaussian noise distributions [12], [35], [36]. This can then be the basis for the integration of 2 or more of the methods. This has been carried out for Wiener and Wavelet filtering in the work of Jin et al. [13]. This work exploited the advantages of both methods simultaneously, but even though the theoretical improvement in peak-to-peak SNR was expected to be 3dB, only 0.5dB was achieved. Also the time complexity of Jin et al's method corresponds to the combination of both the Wiener and wavelet methods. In particular, it would be interesting to evaluate optimised combinations of the Wiener filter and BINR, Wavelet denoising and BINR, and the Wiener-Wavelet and the BINR method, in terms of their computational complexities. These combinations would be expected to yield the advantages of the individual methods, for instance, the smooth background noise presented by the Wiener filter, the reduction in the mean noise power value provided by the BINR method, and the faithful representation of the original signal achieved when applying wavelet denoising.

ACKNOWLEDGMENT

The authors would like to thank NASA/JPL-Caltech for all UAVSAR data.

REFERENCES

[1] J. C. Brailean, R. P. Kleihorst, S. Efstratiadis, A. K. Katsaggelos, and R. L. Lagendijk, "Noise reduction filters for dynamic image sequences: A review," *Proc. IEEE*, vol. 83, no. 9, pp. 1272–1292, Sep. 1995.

[2] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Cambridge, MA, USA: MIT Press, 1964.

[3] B. Widrow *et al.*, "Adaptive noise cancelling: Principles and applications," *Proc. IEEE*, vol. 63, no. 12, pp. 1692–1716, Dec. 1975.

[4] M. Martin-Fernandez, C. Alberola-Lopez, J. Ruiz-Alzola, and C.-F. Westin, "Sequential anisotropic Wiener filtering applied to 3D MRI data," *Magn. Reson. Imag.*, vol. 25, no. 2, pp. 278–292, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0730725X06002852

[5] I. N. Daliakopoulos and I. K. Tsanis, "A weather radar data processing module for storm analysis," *J. Hydroinformat.*, vol. 14, no. 2, pp. 332–344, 2012.

[6] D. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[7] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994. [Online]. Available: http://biomet.oxfordjournals.org/content/81/3/425.abstract

[8] D. L. Donoho and I. M. Johnstone, "Threshold selection for wavelet shrinkage of noisy data," in *Proc. 16th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. Eng. Adv., New Opportunities Biomed. Eng.*, vol. 1. no. 12, Nov. 1994, pp. A24–A25.

[9] C. Burrus, R. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*, vol. 23. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[10] M.-Y. Chen and J.-J. Chao, "Radar image denoising by recursive thresholding," in *Proc. Int. Conf. Image Process.*, vol. 1. Sep. 1996, pp. 395–398.

[11] O. A. M. Aly, A. S. Omar, and A. Z. Elsherbeni, "Detection and localization of RF radar pulses in noise environments using wavelet packet transform and higher order statistics," *Prog. Electromagn. Res.*, vol. 58, pp. 301–317, 2006. [Online]. Available: http://jpier.org/pier/pier.php?paper=0507024

[12] Y. Chen and C. Han, "Adaptive wavelet thresholding for image denoising and compression," *Electron. Lett.*, vol. 41, no. 10, pp. 586–587, May 2005.

[13] F. Jin, P. Fieguth, L. Winger, and E. Jernigan, "Adaptive wiener filtering of noisy images and image sequences," in *Proc. Int. Conf. Image Process. (ICIP)*, vol. 3. Sep. 2003, pp. III-349–III-352.

[14] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 27, no. 2, pp. 113–120, Apr. 1979.

[15] M. I. Skolnik, Ed., "Automatic detection, tracking, and sensor integration," in *Radar Handbook*, 3rd ed. New York, NY, USA: McGraw-Hill, 2008.

[16] P. P. Gandhi and S. A. Kassam, "Optimality of the cell averaging CFAR detector," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1226–1228, Jul. 1994.

[17] A. Foessel, J. Bares, and W. R. L. Whittaker, "Three-dimensional map building with MMW radar," in *Proc. 3rd Int. Conf. Field Service Robot.*, Helsinki, Finland, Jun. 2001. [Online]. Available: http://www.ri.cmu.edu/publication_view.html?pub_id=3758&menu_code=0307

[18] D. Langer, "An integrated MMW radar system for outdoor navigation," in *Proc. IEEE Int. Conf. Robot. Autom.*, Minneapolis, MN, USA, Apr. 1996, pp. 417–422.

[19] M. Barkat, *Signal Detection and Estimation,* ser. Artech House Radar Library. Norwood, MA, USA: Artech House, 2005. [Online]. Available: http://books.google.com/books?id=Le5SAAAAMAAJ

[20] G. Gao, L. Liu, L. Zhao, G. Shi, and G. Kuang, "An adaptive and fast CFAR algorithm based on automatic censoring for target detection in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 6, pp. 1685–1697, Jun. 2009.

[21] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, no. 4, pp. 608–621, Jul. 1983.

[22] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications* (Technology Press Books in Science and Engineering). Cambridge, MA, USA: MIT Press, 1949.

[23] A. Kolmogorov, *Stationary Sequences in Hilbert Space*. Chicago, IL, USA: John Crerar Library, 1978.

[24] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 2, pp. 165–168, Mar. 1980.

[25] D. L. Donoho and I. M. Johnstone, "Ideal denoising in an orthonormal basis chosen from a library of bases," *Comptes Rendus Acad. Sci., Ser. I*, vol. 319, no. 12, pp. 1317–1322, 1994.

[26] G. Luo and D. Zhang, "Wavelet denoising," in *Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology*. Rijeka, Croatia: InTech, 2012, pp. 59–80. [Online]. Available: http://www.intechopen.com/

[27] I. Cohen and B. Berdugo, "Noise estimation by minima controlled recursive averaging for robust speech enhancement," *IEEE Signal Process. Lett.*, vol. 9, no. 1, pp. 12–15, Jan. 2002.

[28] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 4. Apr. 1979, pp. 208–211.

[29] E. Jones, T. Oliphant, and P. Peterson. (2001). *SciPy: Open Source Scientific Tools for Python*. [Online]. Available: http://www.scipy.org/

[30] S. Kuttikkad and R. Chellappa, "Non-Gaussian CFAR techniques for target detection in high resolution SAR images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, vol. 1. Nov. 1994, pp. 910–914.

[31] E. Widzyk-Capehart, G. Brooker, S. Scheding, R. Hennessy, A. Maclean, and C. Lobsey, "Application of millimetre wave radar sensor to environment mapping in surface mining," in *Proc. 9th Int. Conf. Control, Autom., Robot. Vis. (ICARCV)*, Dec. 2006, pp. 1–6.

[32] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

[33] L. Shapiro and G. Sockman, *Computer Vision*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002, ch. 3.

[34] M. di Bisceglie and C. Galdi, "CFAR detection of extended objects in high-resolution SAR images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, vol. 6. Jul. 2001, pp. 2674–2676.

[35] S. Zhong and V. Cherkassky, "Image denoising using wavelet thresholding and model selection," in *Proc. Int. Conf. Image Process.*, vol. 3. 2000, pp. 262–265.

[36] A. Antoniadis, D. Leporini, and J.-C. Pesquet, "Wavelet thresholding for some classes of non–Gaussian noise," *Statist. Neerlandica*, vol. 56, no. 4, pp. 434–453, 2002.

Daniel Lühr (S'99–M'04) was born in Santiago, Chile, in 1978. He received the Lic.Sci. and Diploma degrees in electrical engineering from the University of Chile, Santiago, Chile, in 2003 and 2004, respectively.

He was a Research Associate and No-Fee Consultant with the Submillimeter Receiver Laboratory, Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA, from 2004 to 2006. From 2007 to 2009, he was an Adjunct Professor with the Universidad Austral de Chile, Valdivia, Chile. From 2009 to 2013, he was a Consultant with the Chilean Ministry of Energy. He is currently with the Department of Electrical Engineering, University of Chile. He has expertise in digital systems, system integration, medical devices, and robotics. His current research interests include robotics and its applications.

Martin Adams (SM'08) received the degree in engineering science from the University of Oxford, Oxford, U.K., in 1988, and the Ph.D. degree from the Robotics Research Group, University of Oxford, in 1992. He is currently a Professor of Electrical Engineering with the Department of Electrical Engineering, University of Chile, Santiago, Chile. He is also a Principle Investigator with the industrially sponsored Advanced Mining Technology Centre, University of Chile. After his Ph.D., he continued his research in autonomous robot navigation as a Project Leader and part-time Lecturer with the Institute of Robotics, Swiss Federal Institute of Technology, Zurich, Switzerland. He was a Guest Professor and taught control theory in St. Gallen, Switzerland, from 1994 to 1995. From 1996 to 2000, he served as a Senior Research Scientist in Robotics and Control, in the field of semiconductor assembly automation, with the European Semiconductor Equipment Centre, Cham, Switzerland. From 2000 to 2010, he was an Associate Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research work focuses on autonomous robot navigation, sensing, and multiobject estimation, and has authored many technical papers in these fields. He has been a Principle Investigator and leader of many robotics projects, coordinating researchers from local industries and local and overseas universities, and has served as an Associate Editor of various journal and conference editorial boards.

Appendix B

Underground Mine Dataset paper

Data Paper

The International Journal of
Robotics Research
2017, Vol. 36(1) 16–23

DOI: 10.1177/0278364916679497
journals.sagepub.com/home/ijr

Chilean underground mine dataset

Keith Leung[1], Daniel Lühr[3], Hamidreza Houshiar[4], Felipe Inostroza[2], Dorit Borrmann[4], Martin Adams[2], Andreas Nüchter[4] and Javier Ruiz del Solar[2]

Abstract

This article presents a robotic dataset collected from the largest underground copper mine in the world. The sensor measurements from a 3D scanning lidar, a 2D radar, and stereo cameras were recorded from an approximately two kilometer traverse of a production-active tunnel. The equipment used and the data collection process is discussed in detail, along with the format of the data. This dataset is suitable for research in robotic navigation, as well as simultaneous localization and mapping. The download instructions are available at the following website `http://dataset.amtc.cl`.

Keywords

Field and service robotics, mapping, mining robotics, SLAM, robotics in hazardous fields, sensing and perception, computer vision, sensor fusion

1. Introduction

Located within the slopes of a volcano approximately 100 km south of Santiago, Chile, is the largest underground copper mine in the world. It consists of seven sectors of interconnected mines and we were granted a special permission to enter an active section of the mine, to collect a robotic dataset through a 2 km traverse. Due to the unique opportunity to access an active section of the underground mine with a robotic sensor suite, which has realistic variations in lighting, changes in terrain, elevation, and the presence of water, the dataset could be of interest to researchers in providing robust solutions to robotic navigation and simultaneous localization and mapping (SLAM) in challenging environmental conditions. Many other existing published datasets (Furgale et al., 2012; Peynot et al., 2010; Pomerleau et al.,2012; Tong et al., 2013), and openly accessible data (ASL, 2015; Bonarini et al., 2006; Magnusson et al., 2007; Nüchter and Lingemann, 2015) already provide an abundance of visual and lidar data. However, the particular location where the dataset was collected, along with the availability of the radar data makes this a unique dataset. The purpose of this paper is to describe the method in which the dataset was generated, as well as how it can be interpreted and accessed.

From a robotic mining perspective, this dataset is useful for the preliminary development and evaluation of autonomous vehicles for inspections of mines and caves, abandoned mine exploration, as well as smart mining and material exploitation, where the automated construction of maps is necessary. From a more general robotic navigation and mapping perspective, this dataset is useful for the validation and the benchmarking of sensor fusion and SLAM systems in a realistic environment with variable illumination and surface properties, mainly due to the presence of water, uneven paths, and jagged walls. This dataset could also be used for studying continuous time estimation methods as some of the sensors operated at low-frequency cycles while the robot moved. Streams of information from individual sensors can be used independently to evaluate different types of algorithms such as visual odometry and point cloud registration.

The *Clearpath Robotics Husky A200* platform was used to carry the various on-board sensors, which include a stereo camera, a survey-grade 3D lidar, and a millimeter-wave radar. Figure 1 shows the sensor platform as it traversed a section of the tunnels that was relatively well illuminated. Some parts of the tunnels were dark and had a considerable amount of standing water, as well as water dripping from the ceilings, which makes this dataset challenging to process. For instance, visual-only navigation is

[1]Applanix Corporation (Trimble Inc.), Canada
[2]Department of Electrical Engineering and the Advanced Mining Technology Center (AMTC), Universidad de Chile, Santiago, Chile
[3]Universidad Austral de Chile, Valdivia, Chile
[4]Julius-Maximilians-Universität, Würzburg, Germany

Corresponding author:
Keith Leung, Applanix Corporation (Trimble Inc.), 85 Leek Crescent, Richmond Hill, Ontario, L4B 3B3, Canada.
Email: kleung@applanix.com

Fig. 1. The Husky A200 robot platform carrying a stereo camera, 3D lidar, and a radar, as it traverses through a tunnel in the copper mine for dataset collection.

likely to fail in the parts of the tunnels without lighting, while standing water can reflect lidar beams.

The dataset was collected over a four-hour access window in the tunnels. It is available in both human-readable (text files and png images), as well as in a binary Robot Operating System (ROS) bag format containing standard ROS messages. For organizational purposes, the dataset has been sub-divided into 87 parts. Of the 87 parts, 44 of them were recorded while the robot was stationary at various locations inside the mine. Between the stationary measurement poses, the robot made 43 traverses during which data from various sensors was also recorded. Each of these traverses constitute one part of the dataset. These 43 parts, along with the 44 parts for performing the stationary measurements, make up the entirety of the dataset.

The remainder of this article is structured as follows. Section 2 describes in greater detail the robotic platform and the sensors used in collecting the dataset. Section 3 details the calibration and data collection process. Finally, Section 4 describes the format of the data and ways in which it can be accessed.

2. Equipment and sensors

2.1. Robot platform

The *Clearpath Robotics Husky A200* in Figure 1 was the vehicle used for collecting the dataset. It carried the stereo camera, 3D lidar, and radar on its upper sensor deck, while hosting extra batteries for powering the sensors and two onboard computers on its lower deck. The Husky is a skid-steer vehicle, and both left and right wheel encoder readings were recorded. The wheel-base distance between each pair of front and rear wheels is 50 cm.

Fig. 2. A pair of images from the stereo camera taken at the start of the dataset collection. In relation to Figure 6, the pose of the robot is located at point A, and directed at point B with its forward direction. A gate can be seen in the upper-left corner of both images. The bright orange spots in the upper-center of the images is a pool of water reflecting a set of ceiling-mounted warning lights. The darker spots on the ground are damp surfaces.

2.2. Stereo camera

A *Point Grey XB3* multi-baseline stereo camera was mounted near the front of the Husky A200, with a forward-facing orientation that is also downward-pitched at about 20 degrees to the horizontal plane of motion of the vehicle (see Section 3 for its precise pose). The wider 24 cm baseline configuration for the stereo camera was used. Synchronized left and right (global-shutter) camera images were captured at 16 Hz, both at a resolution of 1280×960 pixels per image. A pair of images from the stereo camera taken from the start of the dataset is shown in Figure 2. Note that in many images, light from headlamps can be observed. Due to safety regulations, they could not be turned off during dataset collection.

2.3. 3D lidar

A *Riegl VZ-400* survey-grade 3D lidar was mounted behind the stereo camera. This sensor has a quoted range limit of 350 m, with a range accuracy of 5 mm, and beam divergence of 0.3 mrad. It is capable of making $125,000$ measurements per second. The generated point cloud contains both positional and intensity information. The intensity information is the relative reflectance in dB, which corresponds to the ratio of the amplitude of the return signal to the amplitude of a white flat (diffuse) target at the same distance and which is aligned orthogonal to the laser beam. Thus, a white target has a 0 dB as intensity (Riegl Laser Measurement Systems, 2009). A relative reflectance higher than 0 dB results from targets reflecting with a directivity different from that of a Lambertian reflector, such as a reflective marker being placed in the environment for surveying. Non-reflections are not reported in this dataset, but the direction (azimuth and elevation) of non-returns could potentially be estimated by first converting a scan into polar coordinates and then searching for the set of azimuth and elevations with the missing data. Scans were obtained at

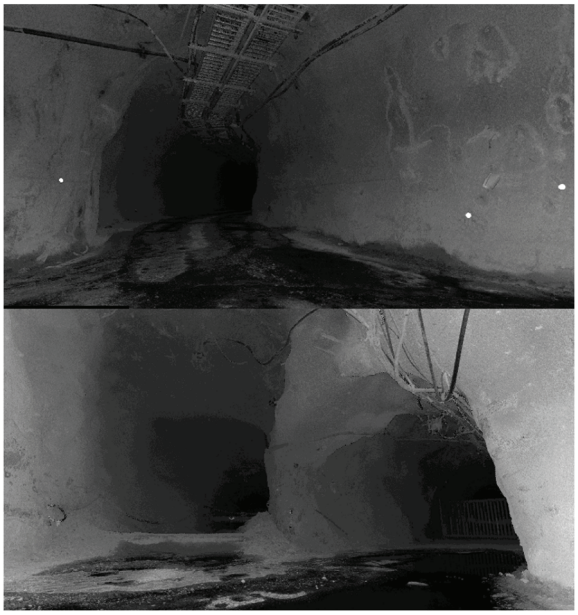

Fig. 3. High resolution point clouds which were logged from the Riegl lidar with high (whiter) pixel values corresponding to high reflectance. Pixel intensities are also attenuated exponentially with respect to the distance to the lidar. In relation to Figure 6, the top scan was collected at Point B, where some reflective markers can be seen (in bright white). These were used for evaluating the accuracy of the pose estimate provided in this dataset. The bottom scan was recorded at point E.

two different resolutions. When the Husky A200 was stationary, high resolution point clouds were recorded at vertical and horizontal resolutions of 0.04 degrees, and at a scan rate of one full scan per 152.5 s^1. When the robot was moving between stationary scan positions, the lidar was commanded to rotate and scan continuously at a higher frequency of one (360 degree) rotation of the sensor every six seconds. This produced scans at lower vertical and horizontal resolutions of 1 degree. Two sample high resolution lidar point clouds which were taken while the robot was stationary are shown in Figure 3.

2.4. 2D radar

An AcuMine frequency-modulated continuous-wave high-speed millimetric (94 GHz) 2D radar (Brooker et al., 2005) was mounted near the rear of the Husky A200. The sensor was operated in a continuous scanning mode where the swashplate rotated at 1 Hz, with a radar sample rate of 435 Hz. Power return values were recorded every 0.5 m up to an operational range of 200 m, and at angular resolutions of 0.68 degrees. The antenna beam width is 1.5 degrees, and the beam is capable of penetrating small dust particles due to its relatively longer wavelength when compared to the lidar. An example of a (cropped) radar scan at a tunnel junction is shown in Figure 4. The origin of the sensor is located

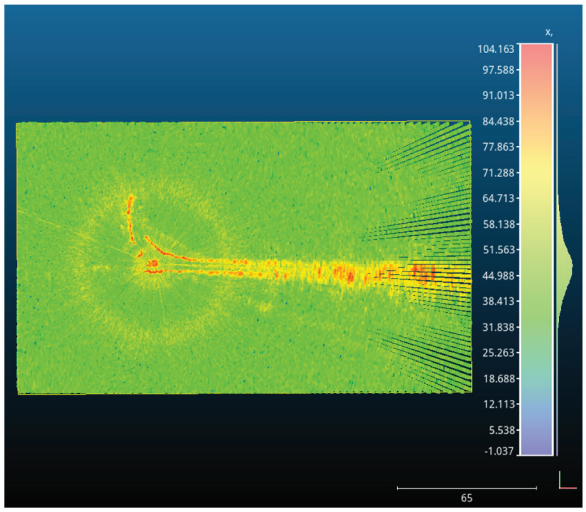

Fig. 4. A 2D radar scan from point B of Figure 6. The sensor is located towards the left side of the figure. The scale for power return values in dB is on the right. Higher power returns (from tunnel walls) are colored in red. A faint yellow halo centered about the sensor is caused by interference from the radar's power supply.

towards the left side of the figure. The faint yellow halo centered about the sensor is an artifact caused by interference from the power supply of the radar. An in-depth explanation of this phenomenon can be found in Adams et al. (2012). At several scan locations, radar data is not available due to hardware problems with the sensor. These locations have been noted in Table 1. Fortunately, due to the range of the sensor, sufficient overlaps exist between the collected radar data to cover the traversed tunnels.

3. Data collection

3.1. Calibration

Prior to the dataset collection, calibration parameters for the stereo camera, as well as the relative pose between each sensor reference frame was determined. These frames are illustrated in Figure 5, with the precise relative transformations reported in Table 2.

The base frame is centered between the four wheels of the Husky A200, with the x-axis aligned with the forward direction. The pose of the lidar frame is fixed by mechanical design. The translation between the lidar and radar frame is also determined by mechanical design, while the relative rotation was calculated by aligning the scans between the two sensors. For the lidar and the stereo camera, distinct tie points were manually selected from multiple sets of scans and images to calculate their relative transformation through least squares optimization. The intrinsic and extrinsic calibration parameters for the stereo camera were obtained by the standard checkerboard calibration process

Table 1. Data description.

Data sequence	Location (see Figure 6)	Notes
00S	Location A	This is the start of the dataset, with the robot facing towards point B. Radar information is not available.
01M, 01S	From A to B	Radar information is not available for 01M.
02M, 02S, . . . , 08S	From B to C	
09M, 09S, . . . , 11S	From C to D	The tunnel between 09S and 10S is dark, but light can be seen from the tunnel ahead.
12M, 12S, . . . , 14S	From D to E	The tunnel between 11S and 14S is dark and turns gradually. Light seen in the stereo camera images are from the safety headlamps of the operators. A large amount of standing and dripping water is present.
15M, 15S	Location E	Ahead of the robot at 15S is a shaft for depositing material that will fall into lower levels of the mine. Radar data is not available.
16M, 16S, . . . , 19S	From E to D	The tunnel is dark. Light seen in the stereo camera images are from the safety headlamps of the operators. A large amount of standing and dripping water is present. Radar data is not available for 16S and 17S.
20M, 20S, 21M, 21S	From D to C	The section traversed in 21M is initially dark. This is the same dark section between 09S and 10S but traversed in the reverse direction.
22M, 22S, . . ., 25S	From C to F	This section is an uphill ascent. The tunnel section between 22S and 24S is dark.
26M, 26S, . . ., 30S	From F to G	Radar data is not available from 28M to 30M.
31M, 31S, . . ., 33S	From G to B	This section is a downhill descent. At the end of 33M, the robot makes a u-turn and climbs slightly uphill towards point G before stopping.
34M, 34S, . . ., 35M	From B to G	This section is an up-hill ascent. Traverse 35M passes through point G without stopping.
35M, 35S, . . ., 37S	From G to F	
38M, 38S, . . ., 40S	From F to C	This section is a downhill descent. Traverse 38M is mostly a dark section of the tunnel that corresponds to the dark section experienced between 22S and 24S. At the end of traverse 40M, the robot turns at junction C to face the direction of junction B.
41M, 41S, . . ., 43S	From C to B	The dataset ends with the robot at point B

(Zhang, 2000). Specifically, the calibration pipeline implementation in OpenCV (Bradski et al., 2000) was used to obtain the camera parameters in Table 3, which assumes the Brown–Conrady distortion model (Brown, 1966).

3.2. Data recording

Two computers were installed on-board the Husky A200 for dataset collection. The first computer, dedicated to data logging, was connected to the Husky's communication port for controlling the motion of the vehicle and receiving wheel encoder measurements. This computer was also connected to the stereo camera and the lidar. The second computer was a dedicated radar server, and all radar measurements were relayed to the first computer for logging.

Throughout the four-hour traverse, the Husky stopped at 44 locations (including the starting position) to record high resolution 3D lidar point clouds, each of about 25 million points, while the robot remained stationary. In straight tunnel sections, the distances between the stopped locations was between approximately 30 m and 40 m to obtain sufficient overlap between the consecutively obtained point

Table 2. The relative transformations between sensor reference frames. The rotation brings a point in frame (b) into frame (a), and the translation is defined from frame (b) to frame (a) with respect to the coordinates defined in frame (a).

Frames (a)	Frames (b)	Rotation (quaternion) $\mathbf{q}_{a,b} = [q_x\ q_y\ q_z\ q_w]$ Translation (m) $\mathbf{p}_a^{b,a} = [p_x\ p_y\ p_z]$
base	lidar	$\mathbf{q}_{a,b} = [0 \quad 0 \quad 1 \quad 0]$ $\mathbf{p}_a^{b,a} = [0.120 \quad -0.003 \quad 0.840]$
lidar	radar	$\mathbf{q}_{a,b} = [0 \quad 0 \quad 0.707 \quad 0.707]$ $\mathbf{p}_a^{b,a} = [0.292 \quad -0.003 \quad 0.401]$
lidar	stereo left	$\mathbf{q}_{a,b} = [0.584 \quad 0.573 \quad -0.407 \quad -0.407]$ $\mathbf{p}_a^{b,a} = [-0.164 \quad -0.106 \quad -0.152]$
stereo left	stereo right	$\mathbf{q}_{a,b} = [0 \quad 0 \quad 0 \quad 1]$ $\mathbf{p}_a^{b,a} = [0.240 \quad 0 \quad 0]$

clouds. These distances were increased when the robot was in a section that it has previously traversed. Radar measurements were also logged when the robot was stationary, but the stereo camera images and the wheel encoder readings were not.

Between the stationary scan poses, 43 traverses were conducted by manually controlling the robot. During each traverse, measurements from the wheel encoders, lidar, radar, and stereo camera were logged. The lidar was configured to scan at a lower resolution to increase the scanning frequency while the robot was moving.

3.3. Time synchronization

A *Network Time Protocol* (NTP) daemon was used to synchronize the clocks between the two on-board computers. The first computer, dedicated to data logging, was setup as the NTP server, while the radar server computer was the NTP client. Sufficient time was allowed after booting up both of the computers so that the client clock could adjust. The average time difference is on the order of 10^{-3} s. Measurements from the wheel encoders, lidar, and stereo camera were time-stamped using the data logging computer's clock, while radar measurements were time-stamped using the radar server's clock.

3.4. Scan pose estimates by supervised pointcloud alignment

The tunnel environment made obtaining the ground-truth robot poses a challenging task due to both occlusions and the aspect ratio of the long and narrow tunnels. Reflective markers were placed at tunnel junctions, but it was infeasible to cover all the traversed tunnels with markers. Since the Riegl is a survey-grade lidar, and the most accurate sensor available on the robot, the pose estimates of the robot at the 44 static scan positions were determined by using the high-resolution 3D point clouds from the lidar. Manual alignment was performed on the consecutive 3D point clouds from the 44 static scanning positions. The batch optimization method in Borrmann et al. (2008) for pose-graph relaxation is then used to refine the alignment[2], and was used to generate Figure 6. The positional errors of the reflective markers, of 50 mm diameter which were observed from various poses, were used to evaluate the quality of the solution, which is a common practice in surveying. The average marker position error between scan pairs is

Table 3. Stereo camera calibration parameters, assuming the Brown–Conrady distortion model (Brown, 1966).

Parameter	Left camera	Right camera
Tangential distortion $\mathbf{p}$	$p_1 = 0.000571 \quad p_2 = 0.000539$	$p_1 = 0.000959 \quad p_2 = 0.000514$
Radial distortion $\mathbf{k}$	$k_1 = -0.177067 \quad k_2 = -0.165155 \quad k_3 = -0.142561$ $k_4 = 0.192250 \quad k_5 = -0.316417 \quad k_6 = -0.196546$	$k_1 = 0.746517 \quad k_2 = -0.321874 \quad k_3 = -0.788330$ $k_4 = 1.099655 \quad k_5 = -0.043691 \quad k_6 = -1.203787$
Rectification matrix $\mathbf{R}$	$\begin{bmatrix} 0.999997 & -0.002331 & -0.001186 \\ 0.002327 & 0.999992 & -0.003369 \\ 0.001194 & 0.003366 & 0.999994 \end{bmatrix}$	$\begin{bmatrix} 0.999993 & -0.003561 & 0.001203 \\ 0.003557 & 0.999988 & 0.003370 \\ -0.001215 & -0.003365 & 0.999994 \end{bmatrix}$
Projection matrix $\mathbf{P}$	$\begin{bmatrix} 997.177459 & 0.000000 & 667.571815 & 0.000000 \\ 0.000000 & 997.177459 & 496.314240 & 0.000000 \\ 0.000000 & 0.000000 & 1.000000 & 0.000000 \end{bmatrix}$	$\begin{bmatrix} 997.177459 & 0.000000 & 667.571815 & -239.608595 \\ 0.000000 & 997.177459 & 496.314240 & 0.000000 \\ 0.000000 & 0.000000 & 1.000000 & 0.000000 \end{bmatrix}$

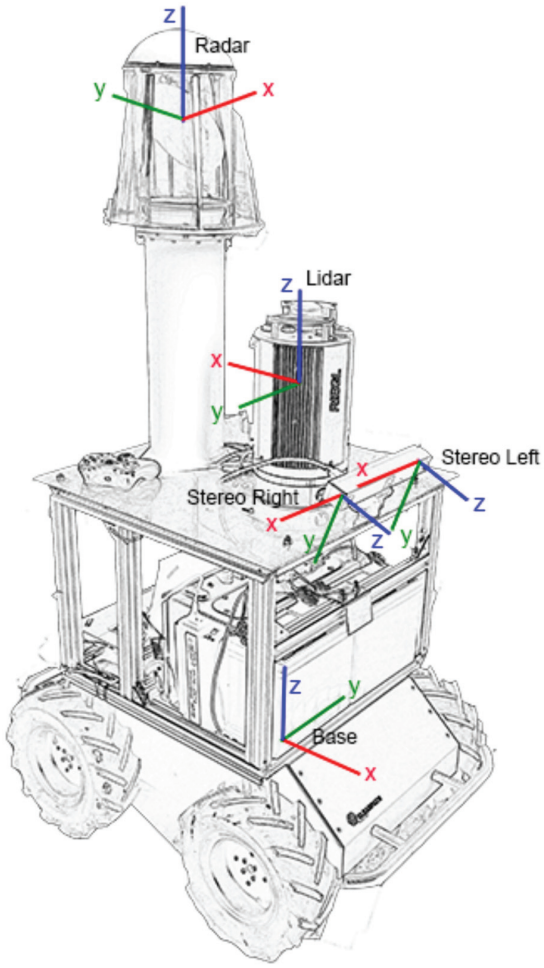

Fig. 5. Sensor reference frames used in the dataset. The base frame in which odometric data is expressed is centered between the four wheels of the Husky. The x, y, and z axes are colored in red, green, and blue, respectively. The precise transformations between the sensor frames are listed in Table 2.

10.41 mm, with the standard deviation being 8.25 mm, and the maximum error being 32.34 mm. The top-view map shown in Figure 6 was constructed based upon all the point clouds being plotted with the 44 pose estimates. Figure 7 shows the alignment between two overlapping point clouds.

3.5. The traverse

With reference to Figure 6, the approximately two-kilometer traverse started at point A, and then proceeded to the tunnel junction at point B. Next, the Husky traversed to the tunnel junction at point C, and continued on to point E through point D. In the vicinity of point D, there was no lighting in the tunnel, and there was a significant amount of standing water, as well as water dripping from the ceiling. Point E is a material drop point, from here the robot turned around and returned to the junction at point C. A right turn was then made to point F, where again there was a lack of lighting. The robot continued to the junction at point B through point G, and then turned around to retrace its path to point C through points G and F. At the junction at point C, the robot turned right, and then returned to the junction at point B.

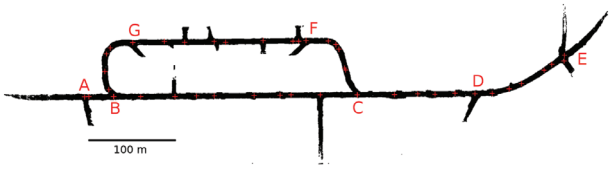

Fig. 6. A view of the map generated from the 3D point clouds at each of the estimated scan poses. Various points of interests are labeled. The data collection traverse initiated at point A and continued to points B, C, D, E, D, C, F, G, B, G, F, C, and finally returned to point B at the end of the dataset. Stationary scan points are marked as crosses.

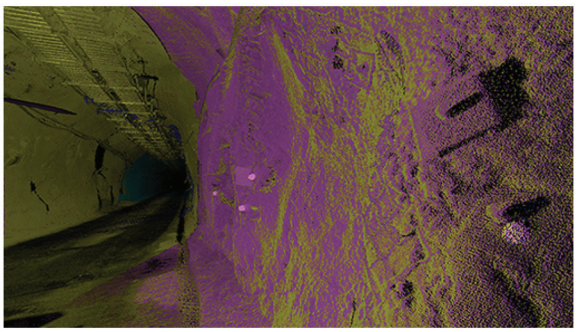

Fig. 7. Two overlapping point clouds shown in different colors, showing the quality of the scan pose estimation method.

4. Data format and usage

The dataset is divided into 87 parts. The data is provided in human-readable text files and png images. The same data is also available as ROS bag files for convenience. For the human-readable files, the following is a description of their file naming convention and their content. Note that each data file also contains a one-line header that describes the columns of the data.

4.1. File naming convention

All of the files with logged sensor data have names that start with a double digit scan index. The letter following the scan index is either an "`S`" or "`M`", and denotes whether the data was gathered when the robot was stationary, or while the robot was moving between stationary scan points, respectively. The data gathered while the robot was stationary are indexed separately from the data gathered while the robot was moving. The data files from the first stationary logging position are indexed with "`00`", while the data from the first traverse are indexed with "`01`". Hence, the dataset was collected in the sequence `00S`, `01M`, `01S`, `02M`, ... , `43M`, `43S`.

The approximate topological position of the robot in relation to Figure 6, along with notes of interest, are provided in Table 1 for each sub-part. Following this alpha-numeric index is a string identifying the type of data that the file contains. This string can be "`_encoder`", "`_stereo`", "`_lidar`", "`_radar`" or "`_timing`". The names of all text files end with "`.dat`", while for stereo image files, the data type string "`_stereo`" is followed by "_", a six digit integer for the image sequence, "`L`" or "`R`" for the left or right camera, and a "`.png`" file extension. The contents of the text files will be described next.

4.2. Wheel encoder file

Each line of an encoder file contains

$$[\text{time}] \quad [\text{enc}_{\text{left}}] \quad [\text{enc}_{\text{right}}]$$

where the time-stamp is in seconds, and enc_{left} and $\text{enc}_{\text{right}}$ are the left and right encoder data, respectively, expressed as distance traveled in meters.

Dead-reckoning estimates from the encoders can be inaccurate due to the skid-steer mechanics of the Husky, as well as the wet, slippery ground surface within the mine tunnels.

4.3. Stereo images

Text files `##X_stereo.dat` contain the time-stamp information for each indexed image pair. Each line in these files is in the following format

$$[\text{time}] \; [\text{index}]$$

where the time-stamp is in seconds and the index is a six-digit integer that corresponds to the index used in the file names of the stereo image `.png` files.

4.4. Lidar point cloud file

Each line in a lidar data file represents one point and is recorded in the format

$$[\text{time}] \quad [x] \quad [y] \quad [z] \quad [i]$$

where the time-stamp is in seconds, followed by the x, y, z 3D coordinate of a point in the lidar reference frame, and its intensity, i, in dB. The center of the lidar is positioned at $(0, 0, 0)$.

4.5. Radar file

A line in a radar data file corresponds to a scan sector (a particular angle of the radar swashplate, and range bin), and has the following format

$$[\text{time}] \quad [x] \quad [y] \quad [p]$$

where the time-stamp is in seconds, followed by x, y, the 2D coordinate representing the center of a received-signal bin (from the discretized measurement space). The last field, p, is the received power in dB.

4.6. Timing file

The timing information from all sensor messages are summarized in the `##X_timing.dat` files. This is to facilitate the reading of sensor messages from their respective data files in the proper order. Each line in a timing file is in the following format

$$[\text{time}] \quad [S]$$

where the time-stamp is in seconds, and S is a letter representing the sensor message. This can be W for the wheel encoder, C for stereo camera, L for lidar, or R for radar.

4.7. Sensor pose file

The relative transformation between the sensor frames of Table 2 are written in the `frames.dat` file. A transformation that brings a point to frame a from frame b using a rotation quaternion $\mathbf{q}_{a,b} = \begin{bmatrix} q_x & q_y & q_z & q_w \end{bmatrix}$ and translation vector $\mathbf{p}_a^{b,a} = \begin{bmatrix} p_x & p_y & p_z \end{bmatrix}$ in units of m is written in two lines in the following format

[frame a]	[frame b]					
$[q_x]$	$[q_y]$	$[q_z]$	$[q_w]$	$[p_x]$	$[p_y]$	$[p_z]$

4.8. Camera calibration file

The file `camera.dat` contains eight lines, with the first four lines containing rectification information for the left stereo camera, and the last four for the right stereo camera. Each set of four lines lists, in order, the tangential distortion parameters, radial distortion parameters, the rectification matrix, and the projection matrix (both in row-major order). The following are the lines corresponding to the left camera

p_l	$[p_1]$	$[p_2]$		
k_l	$[r_1]$	$[r_2]$	...	$[r_6]$
R_l	$[\mathbf{R}_{1,1}]$	$[\mathbf{R}_{1,2}]$	...	$[\mathbf{R}_{3,3}]$
P_l	$[\mathbf{P}_{1,1}]$	$[\mathbf{P}_{1,2}]$	...	$[\mathbf{P}_{4,4}]$

4.9. Estimated pose file

Estimates for the 44 stationary scan poses are recorded in the file `scanPoseEstimates.dat`. Each line begins with the data sequence identification (e.g. 00S for the first stationary data collection pose), followed by its position and orientation (expressed as a rotation quaternion) relative to the first stationary data collection pose

$$[\text{sequence}] \quad [q_x] \quad [q_y] \quad [q_z] \quad [q_w] \quad [p_x] \quad [p_y] \quad [p_z]$$

4.10. Access

Instructions for downloading the dataset are available at the website `http://dataset.amtc.cl`.

Acknowledgments

The authors acknowledge Paul Vallejos and Rodrigo Asenjo from the Advanced Mining Technology Center (AMTC) for facilitating

the logistics that made the dataset collection possible. The authors also gratefully acknowledge the Julius–Maximilians-Universität Würzburg, Germany for their provision of the Riegl 3D lidar.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the Conicyt-DAAD collaborative project PCCI12009 and Fondecyt projects 1150930 and 3150066. The Husky A200 robotic platform was partially supported by the Clearpath Robotics "Partnerbot" program, with the remaining cost of the robot and other sensors being funded by the Universidad de Chile. The support of the AMTC is also acknowledged.

Notes

1. See 2009 Riegl Laser Measurement Systems "Riegl vz-400" data sheet
2. The source code for this method is part of the 3D Toolkit, and is available for download at `http://slam6d.sourceforge.net`.

References

Adams M, Mullane J, Jose E, et al. (2012) *Robotic Navigation and Mapping With Radar*. USA: Artech House.

ASL (2015) ASL datasets repository. Available at: http://projects.asl.ethz.ch/datasets (accessed 24 August 2016).

Bonarini A, Burgard W, Fontana G, et al. (2006) Rawseeds: Robotics advancement through web-publishing of sensorial and elaborated extensive data sets. In: *Proceedings of IROS workshop on benchmarks in robotics research*, Beijing, China, 2006. USA: IEEE.

Borrmann D, Elseberg J, Lingemann K, et al. (2008) Globally consistent 3d mapping with scan matching. *Robotics and Autonomous Systems* 56(2): 130–142.

Bradski G (2000) The opencv library. *Dr. Dobb's Journal of Software Tools* 25(11): 120–126.

Brooker GM, Scheding S, Bishop MV, et al. (2005) Development and application of millimeter wave radar sensors for underground mining. *Sensors Journal* 5(6): 1270–1280.

Brown DC (1966) Decentering distortion of lenses. *Photometric Engineering* 32(3): 444–462.

Furgale PT, Carle P, Enright J, et al. (2012) The Devon island rover navigation dataset. *The International Journal of Robotics Research* 31(6): 707–713.

Magnusson M, Duckett T and Lilienthal AJ (2007) Scan registration for autonomous mining vehicles using 3D-NDT. *Journal of Field Robotics* 24(10): 803–827.

Nüchter A and Lingemann K (2015) 3D robotic scan repository. Available at: http://kos.informatik.uni-osnabrueck.de/3Dscans/ (accessed 24 August 2016).

Peynot T, Scheding S and Terho S (2010) The marulan data sets: Multi-sensor perception in a natural environment with challenging conditions. *The International Journal of Robotics Research* 29(13): 1602–1607.

Pomerleau F, Liu M, Colas F, et al. (2012) Challenging data sets for point cloud registration algorithms. *The International Journal of Robotics Research* 31(14): 1705–1711.

Riegl Laser Measurement Systems (2009) RiSCAN PRO. Software Description & User's Instruction. Version 1-5-0sp1.

Tong C, Gingras D, Larose K, et al. (2013) The Canadian planetary emulation terrain 3d mapping dataset. *The International Journal of Robotics Research* 32(4): 389–395.

Zhang Z (2000) A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 22(11): 1330–1334.

Appendix C

Sensors specifications

C.1 Acumine High Speed Scanning Radar

Parameter	Value
Operational frequency:	94GHz
Transmitted power:	10mW
Antenna beam-width:	1.5°
Angle accuracy:	<0.1°
Operational range:	1 to 200m @ 1rps
Range resolution:	50cm
Rotation rate:	5rps
Measurement output:	Ethernet (Client – Server via TCP)
Supply voltage:	24V DC @ 3A MAX

Table C.1: Radar Operational Specifications

C.2 UAVSAR

Parameter	Value
Operational frequency and band:	1217.5 GHz to 1297.5 GHz(L-band)
Polarization:	Quad-pol
Polarization channels:	HH (red), HV (green), VV (blue)
Nominal bandwidth:	80 MHz
Peak transmitted power:	3.1 kW
Maximum duty cycle:	8 %
Look angle range:	25° to 65°
Nominal range swath:	22 km
Operating altitude range:	2000 m to 18 000 m

Table C.2: Radar Operational Specifications

C.3 Riegl VZ-400 ladar

Technical Data *RIEGL* VZ®-400

Laser Product Classification	Class 1 Laser Product according to IEC60825-1:2007 The following clause applies for instruments delivered into the United States: Complies with 21 CFR 1040.10 and 1040.11 except for deviations pursuant to Laser Notice No. 50, dated June 24, 2007.

Range Performance [1]

	Long Range Mode	High Speed Mode
Laser Pulse Repetition Rate PRR (peak) [2]	100 kHz	300 kHz
Effective Measurement Rate (meas./sec) [2]	42 000	122 000
Max. Measurement Range [3]		
natural targets ρ ≥ 90 %	600 m	350 m
natural targets ρ ≥ 20 %	280 m	160 m
Max. Number of Targets per Pulse	practically unlimited [4]	
Accuracy [5] [7]	5 mm	
Precision [6] [7]	3 mm	
Minimum Range	1.5 m	
Laser Wavelength	near infrared	
Laser Beam Divergence [8]	0.35 mrad	

1) With online waveform processing.
2) Rounded values.
3) Typical values for average conditions. Maximum range is specified for flat targets with size in excess of the laser beam diameter, perpendicular angle of incidence, and for atmospheric visibility of 23 km. In bright sunlight, the max. range is shorter than under overcast sky.
4) Details on request.
5) Accuracy is the degree of conformity of a measured quantity to its actual (true) value.
6) Precision, also called reproducibility or repeatability, is the degree to which further measurements show the same result.
7) One sigma @ 100 m range under *RIEGL* test conditions.
8) Measured at the $1/e^2$ points. 0.35 mrad corresponds to an increase of 35 mm of beam diameter per 100 m distance.

Scanner Performance

	Vertical (Line) Scan	*Horizontal (Frame) Scan*
Scan Angle Range	total 100° (+60° / -40°)	max. 360°
Scanning Mechanism	rotating multi-facet mirror	rotating head
Scan Speed	3 lines/sec to 120 lines/sec	0°/sec to 60°/sec [10]
Angular Step Width Δ ϑ (vertical), Δ φ (horizontal)	0.0024° ≤ Δ ϑ ≤ 0.288° [9] between consecutive laser shots	0.0024° ≤ Δ φ ≤ 0.5° [9] between consecutive scan lines
Angle Measurement Resolution	better 0.0005° (1.8 arcsec)	better 0.0005° (1.8 arcsec)
Inclination Sensors	integrated, for vertical scanner setup position, details see page 2	
GPS Receiver	integrated, L1, with antenna	
Compass	optional, for vertical scanner setup position, details see page 2	
Internal Sync Timer	integrated, for real-time synchronized time stamping of scan data	
Scan Sync (optional)	scanner rotation synchronization	

9) Selectable.
10) Frame scan can be disabled, providing 2D scanner operation.

General Technical Data

Power Supply Input Voltage	11 - 32 V DC
Power Consumption	typ. 65 W (max. 80 W)
External Power Supply	up to three independent external power sources can be connected for uninterrupted operation
Main Dimensions	Ø 180 x 308 mm (diameter x length)
Weight	approx. 9.6 kg
Humidity	max. 80 % non condensing @ +31°C
Protection Class	IP64, dust- and splash-proof
Temperature Range	
Storage	-10°C up to +50°C
Operation	0°C up to +40°C: standard operation
Low Temperature Operation [11]	-20°C: continuous scanning operation if instrument is powered on while internal temperature is at or above 0°C and still air -40°C: scanning operation for about 20 minutes if instrument is powered on while internal temperature is at or above 15°C and still air

11) Insulating the scanner with appropriate material will enable operation at even lower temperatures.

RIEGL® LASER MEASUREMENT SYSTEMS

RIEGL Laser Measurement Systems GmbH
Riedenburgstraße 48
3580 Horn, Austria
Phone: +43 2982 4211 | Fax: +43 2982 4210
office@riegl.co.at
www.riegl.com

RIEGL USA Inc.
Orlando, Florida | info@rieglusa.com | www.rieglusa.com

RIEGL Japan Ltd.
Tokyo, Japan | info@riegl-japan.co.jp | www.riegl-japan.co.jp

RIEGL China Ltd.
Beijing, China | info@riegl.cn | www.riegl.cn

www.riegl.com

Appendix D

Normalized OSPA metric

In some situations, it might be desirable to compare OSPA results with different cutoff parameters. A dimension-less distance based on OSPA can be used for such cases. The following proof shows that a normalized OSPA $\bar{d}_n^{(v)}$ measurement calculated as the OSPA distance $\bar{d}^{(v)}$ divided by the OSPA cutoff parameter v is still also a valid metric.

$$\bar{d}_n^{(v)} = \frac{\bar{d}^{(v)}}{v} \tag{D.1}$$

The normalized OSPA satisfies all metric space conditions:

1. Separation axiom:

$$\bar{d}_n^{(v)} \geq 0 \tag{D.2}$$

$$\bar{d}_n^{(v)} = \frac{\bar{d}^{(v)}}{v} \geq 0 \tag{D.3}$$

   The OSPA metric satisfies the separation axiom, therefore by dividing it by the cutoff parameter, which is always positive, the normalized OSPA is also non-negative.

2. Identity of indiscernibles:

$$\bar{d}_n^{(v)}(\mathcal{A}, \mathcal{B}) = 0 \Leftrightarrow \mathcal{A} = \mathcal{B} \tag{D.4}$$

$$\bar{d}_n^{(v)}(\mathcal{A}, \mathcal{B}) = \frac{\bar{d}^{(v)}(\mathcal{A}, \mathcal{B})}{v} \tag{D.5}$$

   if $\mathcal{A} = \mathcal{B} \Leftrightarrow \bar{d}^{(v)}(\mathcal{A}, \mathcal{B}) = 0$, thus $\bar{d}_n^{(v)}(\mathcal{A}, \mathcal{B}) = 0$.

3. Symmetry:

$$\bar{d}_n^{(v)}(\mathcal{A},\mathcal{B}) = \bar{d}_n^{(v)}(\mathcal{B},\mathcal{A}) \tag{D.6}$$

$$\frac{\bar{d}^{(v)}(\mathcal{A},\mathcal{B})}{v} = \frac{\bar{d}^{(v)}(\mathcal{B},\mathcal{A})}{v} \tag{D.7}$$

The OSPA metric is symmetric, then the last equation is true.

4. Triangle inequality:

$$\bar{d}_n^{(v)}(\mathcal{A},\mathcal{C}) \leq \bar{d}_n^{(v)}(\mathcal{A},\mathcal{B}) + \bar{d}_n^{(v)}(\mathcal{B},\mathcal{C}) \tag{D.8}$$

$$\frac{\bar{d}^{(v)}(\mathcal{A},\mathcal{C})}{v} \leq \frac{\bar{d}^{(v)}(\mathcal{A},\mathcal{B})}{v} + \frac{\bar{d}^{(v)}(\mathcal{B},\mathcal{C})}{v} \quad v \geq 0, \text{ multiplying by } v \text{ at both sides} \tag{D.9}$$

$$\bar{d}^{(v)}(\mathcal{A},\mathcal{C}) \leq \bar{d}^{(v)}(\mathcal{A},\mathcal{B}) + \bar{d}^{(v)}(\mathcal{B},\mathcal{C}) \tag{D.10}$$

Equation (D.10) holds, because the OSPA metric itself satisfies the triangle inequality.

Therefore, the normalized OSPA is also a metric. For other metric normalization examples see [Yia02].