

Segmentation and Classification for Through-the-Wall Radar Imaging

Vom Fachbereich 18
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

von
Ahmed A. Mostafa, M.Sc.
geboren am 28.01.1976 in Alexandria, Ägypten.

Referent:	Prof. Dr.-Ing. Abdelhak M. Zoubir
Korreferent:	Prof. Dr. Wilhelm Stannat
Tag der Einreichung:	02.07.2012
Tag der mündlichen Prüfung:	05.11.2012

Acknowledgments

This research project would not have been possible without the support of many people. I wish to thank everyone who have helped and inspired me during my doctoral study.

The author wishes to express his gratitude to Prof. Dr.-Ing. Abdelhak Zoubir for being an older brother before being a supervisor. His kindness and wisdom combined with being an outstanding researcher really motivated me throughout my PhD study.

I wish to thank Prof. Dr. Wilhelm Stannat for his supervision and guidance. I benefited greatly from our interactions.

Special thanks to Prof.Moeness Amin and Prof. Fauzia Ahmad from CAC at Villanova University for providing the images database.

Deepest gratitude are also due to the members of the supervisory committee, Prof. Dr.-Ing. Ralf Steinmetz, Prof. Dr.-Ing. Franko Küppers and Prof. Dr. rer. nat. Andy Schürr who acted as examiners and chair in the PhD defense committee.

I wish to especially thank my friend Weaam Alkhaldi, who directed me to this opportunity, I will always be grateful. Special gratitude goes to Dr.Christian Debes and Gökhan Gül for their participation and help in my journals. My office room mates and my friends Waqas Sharif, Fiky Suratman, Arezki Younsi and Toufik Boukabba, you really made my days. My thanks go to my colleagues at the Signal Processing Group at TU Darmstadt. I was, and still am, very happy to work in such a sociable and friendly environment. Thanks to Stefan Leier, Michael Muma, Yacine Chakhchoukh, Zihua Lu, Nevine Demitry, Sara Al-Sayed, Mouhammad Alhumaidi, Michael Fauss, Jürgen Hahn, Michael Leigsnering, Feng Yin, Christian Weiss and Gebremichael Teame, as well as Renate Koschella and Hauke Fath. I would also like to thank the former PhD students Raquel Fandos, Uli Hammes, and Marco Moebus. Special thanks go to my friends Ahmed Abdelkader, Mohamed Morsy, Mohamed Said, Wael Said, and Mohamed Youssef for making my stay here in Germany more fruitful and easier.

The author would also like to convey thanks to the Graduate school for computational engineering at TU-Darmstadt for providing the financial means and facilities and its remarkable staff for their help during my stay at Germany.

I wish to thank my parents Abdelsalam Mostafa & Samia Hassaan and my Aunt Thuraaya for their unconditional love, support and patience throughout my life, whatever I say will never be enough. I would also like to thank the rest of my family, especially Zeyad, Yasmin, Mohamed, Shady, Asser, Enas, Eman and Dina for believing in me more than myself.

A heart-felt gratitude to my wife Eman and my sons Abdelrahman and Youssef for their love, encouragement, understanding, support and joy, you are the start and the end, I dedicate this work for you.

Finally, before and above all, I thank GOD for all the blessings I have, and for answering my prayers and giving me the faith and will to continue despite all the obstacles and the hard times I faced during this thesis project.

Darmstadt, 25.05.2012

Kurzfassung

In dieser Doktorarbeit wird das Problem der stationären Zielerfassung, Segmentierung und Klassifikation in Through-the-Wall Radar Imaging (TWRI) betrachtet. In der stationären Zielerfassung sind Techniken, die auf dem Doppler effekt oder auf Veränderungen in der Szene basieren, nicht anwendbar. Eine neue Menge von Merkmalen wird verwendet, die von polarimetrischen Signaturen und so genannter *co-occurrence* Matrizen, die gleichzeitiges Auftreten bestimmter Eigenschaften reflektieren, abhängig ist. Algorithmen für die 2D- und 3D-Segmentierung und Klassifikation werden angepasst, untersucht und getestet.

Die Anwendung dieser Algorithmen in TWRI wird mit besonderem Fokus auf die Merkmalsextraktion und Klassifikation der Ziele untersucht. Eine Kombination aus polarimetrischen Signaturen und Merkmalen, die aus den *co-occurrence* Matrizen extrahiert wurden, werden vorgeschlagen.

Zwei unterschiedliche Systeme, die für 2D- und 3D Anordnungen ausgelegt sind, werden präsentiert. Die erste Methode basiert auf einer Verschmelzung der zweidimensionalen Segmentierung und der Klassifikation. Hier werden Merkmale von polarimetrischen B-Scans zur Segmentierung und Klassifikation des beobachteten Bildes in Ziel, Stördaten und Rauschen verwendet. Polarimetrische Signaturen des Zieles von kopolarisierten und kreuzpolarisierten Zielechos werden Pixel für Pixel in den Merkmalsraum abgebildet. Das Bild wird hierdurch über-segmentiert, um homogene Regionen, so genannte Superpixel, zu erhalten. Homogene Superpixel werden wahlweise im Anschluss gruppiert und dann entsprechenden Klassen zugewiesen. Diese Methode beruht auf neuen Features und den Beziehungen zwischen den verschiedenen Bildpunkten.

Die zweite Methode befasst sich direkt mit der 3D-Szene anstatt der zweidimensionalen B-Scans. Im ersten Schritt werden Gruppierungen nach Intensität und räumlichen Merkmalen für jedes Voxel durchgeführt. Gruppen, die hauptsächlich Rauschen beinhalten, werden ausgeschlossen. Aus den Voxeln der verbleibenden Gruppen werden weitere Merkmale extrahiert, die aus der *co-occurrence* Matrix und der polarimetrischen Signaturen erzeugt werden. Anschließend werden die Voxel mit verschiedenen Klassifikatoren klassifiziert, um die Nützlichkeit der Features zu prüfen. Diese Methode wurde für praktische Anwendungen, bei denen die Erfassung der Ziele in Echtzeit durchgeführt werden soll, entworfen. Die Gruppierung im ersten Schritt wird verwendet, um die Positionen der Ziele schnell zu erkennen. Weitere Schritte werden angewandt, um genauere Schätzungen der Zielpositionen

und Formen zu erhalten und die gruppierten Voxel mit einer höheren Genauigkeit klassifizieren zu können.

Alle vorgeschlagenen Methoden werden anhand realer Daten ausgewertet. Die Daten wurden durch dreidimensionale Messungen eines Breitband-Radar-Bildgebungssystems mittels Beamforming erfasst.

Abstract

In this thesis, the problem of stationary target detection, segmentation and classification in Through-the-Wall Radar Imaging (TWRI) is considered. In stationary target detection, Doppler and change-detection-based techniques are inapplicable. A new feature-set that depends on polarimetric signatures and co-occurrence matrices are employed. Algorithms for 2D and 3D segmentation and classification are adapted, investigated and tested.

The utilization of these algorithms to the application of TWRI is investigated, with special focus on the feature extraction and target classification phases. A combination of polarimetric signatures and features extracted from co-occurrence matrices is proposed.

Two different schemes that deal with 2D and 3D arrangements are proposed. The first scheme is based on a fusion of two dimensional segmentation and classification. The proposed scheme uses features from polarimetric B-scan images to segment and classify the image observations into target, clutter, and noise segments. Target polarimetric signatures from co-polarized and cross-polarized target returns are mapped to a pixel-by-pixel feature space. The image is then over-segmented to homogeneous regions called super-pixels. Homogeneous super-pixels are optionally grouped into clusters and then assigned to corresponding classes. This scheme relies on novel features and the relations between the different pixels.

The second scheme deals with the 3D scene directly instead of 2D B-Scans. The proposed scheme uses clustering as an initial phase using intensity and spatial features for each voxel. Clusters that contain mostly noise are ruled out. Further feature extraction using features from the multivariate co-occurrence matrices and polarimetric signatures is applied to the voxels of the remaining cluster(s). Subsequently, the voxels are classified using different classifiers to test the usefulness of the features. This method is designed for practical applications in which the target detection should be performed in real time. The clustering step is used to detect the target positions quickly. Further steps are used to obtain more accurate estimates of the target positions and shapes, and further classify the clustered voxels.

All proposed methods are evaluated using real data measurements. The data are collected using three dimensional imaging measurements in a wideband radar imaging scanner exploiting wideband delay-and-sum beamforming.

Contents

1	Introduction and Motivation	1
1.1	Motivation	1
1.2	State-of-the-Art	2
1.3	Contributions	3
1.4	Publications	4
1.5	Thesis overview	4
2	Image formation and Experimental Setup	5
2.1	Beamforming in Through-the-Wall Radar Imaging	5
2.2	Experimental Setup	8
3	Feature-sets	11
3.1	Motivation	11
3.2	Polarization Signatures Sets	11
3.3	Co-occurrence matrix	15
3.3.1	Multivariate Co-occurrence Matrices	17
3.3.2	Texture features from the multi-variate co-occurrence matrices	19
3.4	Conclusion	22
4	Segmentation by classification	23
4.1	Motivation	23
4.2	Segmentation	24
4.3	Clustering	26
4.4	Classification with Random Forests	27
4.5	Experimental Results	32
4.5.1	Experiments description	32
4.5.2	Target, noise and clutter classification	34
4.5.3	Further classification, metal/non-metal and target type clas- sification	39
4.5.4	General results	42
4.6	Conclusions	45
5	Multi-level 3D classification	49
5.1	Motivation	50
5.2	Feature Extraction	51
5.3	Clustering	52
5.4	Classification	53

5.4.1	Classification using the estimated probabilities of co-occurrence matrices (EPCM)	55
5.4.1.1	Target / non-target classification	55
5.4.1.2	Multi-class Classification	59
5.4.2	Classification Using Polarimetric Signatures and Texture Features (PolTex)	60
5.5	Experimental results	60
5.5.1	Classification experiments using the co-occurrence matrix estimated probabilities	61
5.5.1.1	Experiments description	61
5.5.1.2	Experiments results: optimum gray level	61
5.5.1.3	Experiments results: classification	62
5.5.2	Classification experiments using polarimetric signatures and texture features (PolTex)	65
5.5.2.1	Experiments description	65
5.5.2.2	Experiments results	65
5.6	Conclusions	68
6	Conclusions and Future Work	71
6.1	Conclusions	71
6.1.1	Features	71
6.1.2	Segmentation by Classification	72
6.1.3	Multi-level 3D Classification	72
6.2	Future Work	73
6.2.1	Super-voxels	73
6.2.2	Clutter and noise reduction	73
6.2.3	Features	74
6.2.4	Classification	74
6.2.5	Wall Removal	75
	List of Acronyms	77
	List of Symbols	79
	Bibliography	81
	Curriculum vitae	95

Chapter 1

Introduction and Motivation

Through-the-Wall Radar Imaging (TWRI) is a research field of increasing interest [1–6]. The ability to sense through visually opaque material and man made structures, has numerous applications. Applications in civil engineering, search and rescue operations, cultural heritage diagnostics, law enforcement, and military applications [1, 3, 7–9] make it a highly desirable tool. TWRI can be used in military and homeland security purposes to detect and classify concealed weapons and explosives [10, 11]. It can also detect buried people after natural (or man-made) disasters, e.g. earthquakes, or after a bomb or missile hit. In hostage crisis situations, it might allow police units to detect and locate hostages, hostage-takers and weapons before entering the building. In all these applications, radio frequency (RF) emission and reception are exploited to gain vision into scenes that are inaccessible physically, optically, acoustically, or thermally. TWRI is faced with many challenges regarding detection and classification. Among these challenges, a large variety of possible indoor targets may exist in the presence of multi-path and unwanted wall signal attenuation and dispersive effects [8, 9, 12, 13]. In addition, Doppler signatures and change detection techniques become ineffective when considering stationary targets. One has to perform detection and classification in the image domain, as a post-processing step to beamforming.

Robust computer-based systems that use automatic target detection and classification frameworks are of high practical interest in this area. This thesis aims to introduce new TWRI target classification methods that exploit new feature-sets and need no or only marginal prior knowledge on scene statistics.

1.1 Motivation

TWRI is an evolving technology that gained much attention in the last decade. Tasks like automatic detection of objects of interest, target classification and tracking, image interpretation and understanding are of high practical interest [3, 8]. The variance of target size, shape, electric properties, and surroundings, among other things, highly influence the image statistics. With several possible indoor targets such as human, furniture, and appliances, as well as the influence of wall impairing

and multipath propagation effects, robust detection in which the detector adapts itself to the changing and unknown characteristics of the data is crucial [11]. Targets with different geometrical and physical properties generate different polarizations of the scattered field. Since the polarization signatures are different, they can be used to discriminate between different classes in TWRI scenes [10, 14]. A feature-set that will depend on a combination of polarimetric signatures can enhance the performance of detection and classification techniques [15–21]. The target pixels and the closely surrounding noise and clutter pixels usually exhibit small variations between each other which make the target hard to detect. Algorithms that measure texture are built to deal with such small variations even in the case of non-optical images like SAR and MRI images [22]. The Gray level co-occurrence matrix is one of the most well-known texture features [23–25], containing texture, histogram and edge information [26]. The use of features extracted from co-occurrence matrices can aid in the discrimination between target and the closely surrounding non-target pixels. Incorporating the information provided by the third dimension will also enhance the overall performance. Jointly or sequentially combining these two feature-sets may further enhance the performance of target classification algorithms. Practical methods should take into account the time and computational complexity as a critical factor regarding the nature of TWRI applications. Often a better result can be obtained by a reduced (optimized) feature set relieving also computational load. Techniques that determine and exploit such an optimized set of features is of great practical interest. Also, Multi-level algorithms that will produce preliminary information about targets then allow for fine tuning of this information are highly desirable.

1.2 State-of-the-Art

Due to the complexity of TWRI, different cross-disciplinary research areas are investigated. In the area of electromagnetic propagation and modeling, the factors governing the signal propagation in the indoor environment and in different material mediums are examined [12, 27–29]. Estimation of the location and physical parameters of the targets by solving the inverse problem after estimating the initial conditions is investigated in [9, 30, 31] among others. In [32–36], the antenna and waveform design is analyzed and optimized to achieve high quality images. Beamforming and image formation is a main area of research in TWRI [2, 31, 37–46]. In [2, 31, 40, 41, 47–50], wall compensation considering the effect of wave propagation and wall reflections is investigated in [9, 51, 52]. Preprocessing in the sense of clutter removal and noise reduction is investigated in [53–55]. The use of compressive

sensing to reduce or use fewer measurements is considered in [56–59]. In [60–64] among others, general image processing techniques such as interpretation, alignment, and reconstruction to produce high quality images and test quality measures are investigated. Most contributions in the area of target detection in TWRI deal with detecting and tracking moving targets, where Doppler shifts and change detection can be employed [12, 65–71]. In stationary target detection, in the image domain, most of the work is done by Debes et al. [72–80]. Other work in stationary target detection in the image domain includes [2, 81–83]. On target classification, there is the work of Debes et al. [11, 84, 85] where resolution independent statistical and geometrical features are used for target classification, but images need to be segmented and classified to (target /non-target) classes first. Rosenbaum and Mobasserri [86] also investigated target classification for TWRI where the principal component analysis (PCA) is applied. However, this approach is practically limited as the authors provide features which are not resolution-independent [11].

1.3 Contributions

The contributions of this thesis are as follows:

- **New Feature-set:** Two feature-sets are introduced for TWRI imaging applications. The first depends on a combination of polarization features considering both values for each pixel or voxel by itself or spatially concerning the neighborhood of each voxel. The second feature-set makes use of the estimated probabilities or relative frequencies from the co-occurrence matrices and texture features extracted from them.
- **Segmentation by classification:** A target classification framework that is based on polarimetric signatures, super-pixels and decision trees is introduced [87]. The algorithm classifies between target /non-target, target types and metal / non-metal classes in different phases. It proves comparable performance to the current state-of-the-art techniques in target detection and classification.
- **3D Target Classification:** A 3D target classification framework, consisting of feature extraction, clustering and classification by exploiting polarimetric signatures and co-occurrence matrices is formulated and adapted to the application of TWRI. Here, the features incorporate three dimensional information and relations between voxels.

1.4 Publications

The following publications have been produced during the period of PhD candidacy.

Internationally Refereed Journal Articles

- Ahmed A. Mostafa, C. Debes and A.M. Zoubir Segmentation by Classification for Through-the-Wall Radar Imaging Using Polarization Signatures, IEEE Transactions on Geoscience and Remote Sensing, September 2012.

Internationally Refereed Conference Papers

- Ahmed A. Mostafa and A. M. Zoubir, 3D Target Detection for Through the Wall Radar Imaging, SPIE Defense, Security, and Sensing Volume 7697 of Signal Processing, Sensor Fusion, and Target Recognition XIX, page 76971F-76971F-8. 2010.

1.5 Thesis overview

The outline of the thesis is as follows: Chapter 2 describes the image formation and beamforming process used to obtain three-dimensional TWRI images. It further introduces the experimental setup that is used throughout the thesis.

Chapter 3 introduces the feature-sets used for target detection, segmentation and classification.

In Chapter 4, a segmentation by classification framework is introduced which makes use of polarimetric signatures to find the homogeneity between neighboring pixels helping to generate super-pixels. The super-pixels are further classified to different classes by means of decision trees.

3D target classification is considered in Chapter 5. A classification framework that exploits the spatial features between voxels through polarimetric signatures and co-occurrence matrices is presented. Further, these features are used to classify the voxels to different classes.

Conclusions and future work are presented in Chapter 6.

Chapter 2

Image formation and Experimental Setup

Image formation and experimental setup for TWRI that is used throughout the thesis are discussed in this chapter. The radar images used throughout this thesis are generated using a wideband delay-and-sum beamformer (DSBF) [2, 41]. For TWRI data examples treated in this paper, perfect knowledge or correctly estimated values of the wall parameters are assumed. Estimation techniques of the wall thickness and dielectric constant can be found in [4, 5] and references therein.

2.1 Beamforming in Through-the-Wall Radar Imaging

The signals received by antenna arrays need to be processed in order to obtain an intensity image. This process is called 'beamforming'. Many approaches exist for beamforming for TWRI applications [11, 88]. This includes tomographic approaches [31, 37, 38] where the beamforming is considered as an inverse scattering problem, differential SAR [52], constrained minimization radio frequency (RF) multisensor processing as in minimum variance distortion-less response (MVDR) beamforming [46], MIMO processing [45] and adaptive beamformers [41, 89]. In this thesis, wideband delay-and-sum beamforming [2] is used. All algorithms mentioned in the following chapters are independent of the beamforming method and can be applied to output images of any beamforming algorithm.

We hereby follow the same scheme as in [2]. We consider a uniform array of N transceivers using a wideband pulse, approximated by a stepped frequency approach. Stepped frequency implies M monochromatic signals with regularly spaced frequencies that cover the desired bandwidth. Assuming that the scene of interest can be described by P discrete targets with reflectivity σ_p , $p = 0, \dots, P - 1$, the received signal, which is a function of the sensor element n , $n = 0, \dots, N - 1$ and frequency m , $m = 0, \dots, M - 1$, can be written as

$$y[m, n] = \sum_{p=0}^{P-1} \sigma_p w_m e^{-j2\pi f_m \tau_{pn}} \quad (2.1)$$

where f_m denotes the m -th step frequency, τ_{pn} is the round trip delay between the p -th target and the n -th receiver and w_m is the weighting factor for the m -th frequency.

A derivation can be found in [2, 90]. Assuming the region of interest is identified by $Q = N_x \cdot N_y \cdot N_z$ voxels, where N_x is the number of points in crossrange, N_y is the number of points in downrange and N_z is the number of points in height (number of B-Scans), one can steer the beam at every voxel q with $q = 0, \dots, Q - 1$ and the complex valued scene can be obtained by [2]:

$$I(q) = \frac{1}{MN} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} y[m, n] e^{j2\pi f_m \tau_{qn}} \quad (2.2)$$

where τ_{qn} is the delay compensation for the n -th receiver (incorporating the propagation through the wall), steering the beam at position q . At a certain height h , the delay compensation τ_{qn} can be represented on a local coordinate system (u, v) where $q = (u_q, v_q, h)$ as:

$$\tau_{n,\text{wall}}(q) = (R_{n,\text{air},1}(q) + \sqrt{\epsilon} R_{n,\text{wall}}(q) + R_{n,\text{air},2}(q))/c \quad (2.3)$$

where ϵ is the dielectric constant of the wall and $R_{n,\text{air},1}(q)$, $R_{n,\text{wall}}(q)$ and $R_{n,\text{air},2}(q)$ represent the traveling distances of the electromagnetic wave from the n -th antenna to point (q) before, through and beyond the wall respectively [11] as shown in Figure 2.1.

The distances $R_{n,\text{air},1}(q)$, $R_{n,\text{wall}}(q)$ and $R_{n,\text{air},2}(q)$ can be estimated as in [2]

$$R_{n,\text{air},1}(q) = \frac{u_{\text{off}}}{\cos(\varphi_{n,I}(q))} \quad (2.4)$$

$$R_{n,\text{wall}}(q) = \frac{d}{\cos(\varphi_{n,R}(q))} \quad (2.5)$$

$$R_{n,\text{air},2}(q) = \frac{u_q - u_{\text{off}} - d}{\cos(\varphi_{n,I}(q))} \quad (2.6)$$

where u_{off} is the standoff distance from the system to the wall, u_q is the down range distance from the system to the imaged point (q) , d is the wall thickness, and $\varphi_{n,I}(q)$ and $\varphi_{n,R}(q)$ are the angles of incidence and refraction, respectively. It should be noted that the above equations hold only when the imaged point (q) and the transceiver are at the same height h . For the general case, a rotation transformation as in [2] has to be performed.

Samples of B-Scan images at the height of the target centers are shown in Figure 2.2.

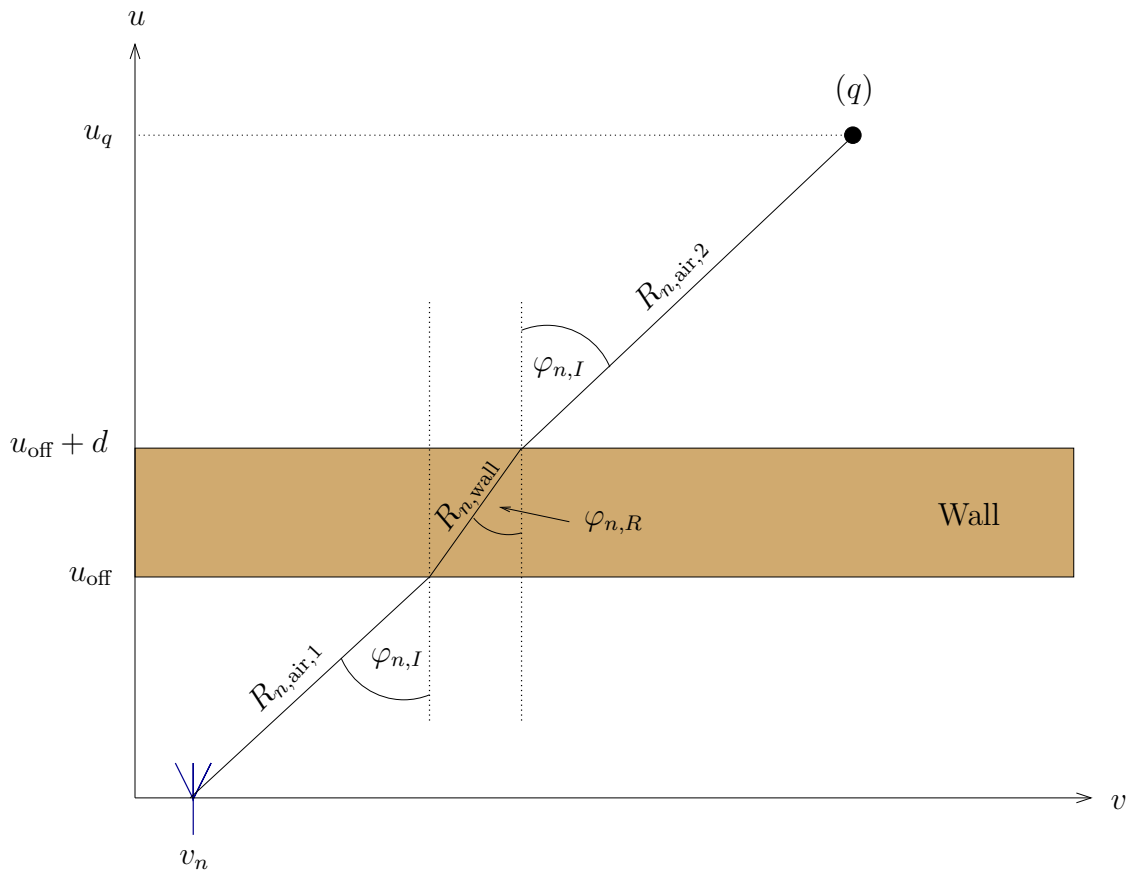


Figure 2.1. Propagation before, through and beyond a homogeneous wall

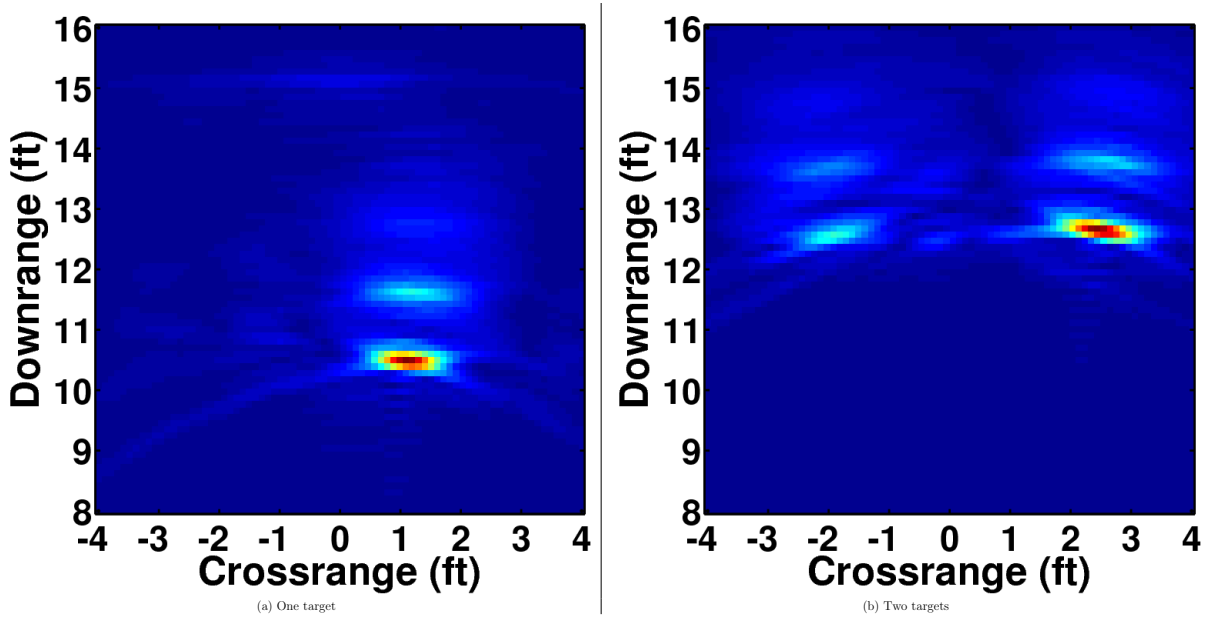


Figure 2.2. Sample beamformed images using DSBF beamforming.

2.2 Experimental Setup

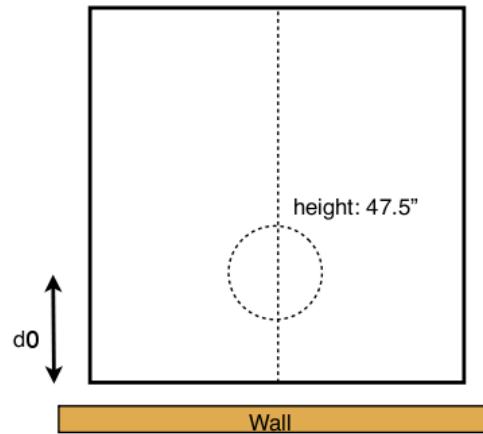
We consider the scenario presented in Section 2.1 for the evaluation of the proposed techniques. The imaging system used throughout this thesis is a synthetic aperture radar system [91], where a single horn antenna, in motion, synthesizes a 57×57 element planar array. The inter-element spacing is $0.875''$. The array standoff distance from the wall is $41.5''$. As described above, a continuous-wave stepped-frequency signal is used to approximate a wideband pulse. The background subtraction technique [2] is used to increase the signal to clutter power ratio. Background subtraction has been performed by making use of reference or background data (in this case: a room without objects) and coherent subtraction. This reference data may be secured in long-term surveillance operations where new targets emerge over time.



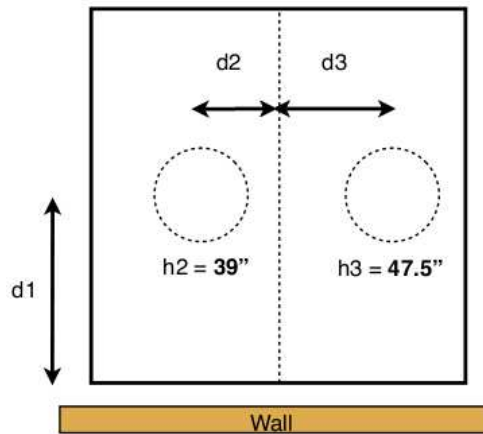
Figure 2.3. Indoor scene for an experiment with one dihedral at the radar imaging lab at Villanova university, USA.

A photo of the indoor scene with a dihedral in the radar imaging lab is shown at Figure 2.3. The experimental setup is depicted in Figure 2.4, which involves seven experiments. The first four experiments involve one target from the following: metal dihedral, metal sphere, metal trihedral and a salt water jug placed at a height of $47.5''$. The fifth experiment involves two dihedrals and the sixth experiment includes a metal sphere and a salt water jug. The heights of the objects on the fifth and sixth experiments are $39''$ and $47.5''$. The last experiment contains four targets: metal dihedral, metal sphere, metal trihedral and a salt water jug placed at heights of $69.5''$, $47.5''$, $63.5''$ and $39''$ respectively. The targets are placed on a high foam

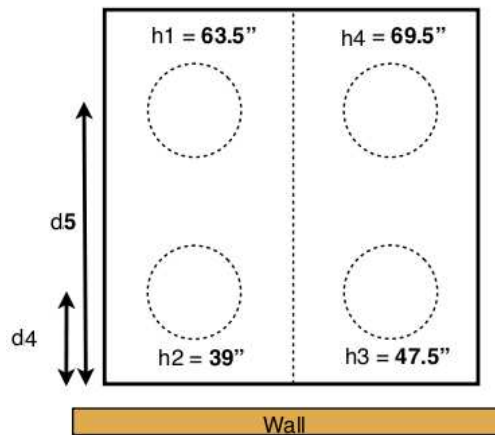
column behind a concrete wall of thickness 5.622". All the experiments are taken from a single side. All images are acquired at the Radar Imaging Lab at Villanova University, Villanova, PA, USA.



(a) Experiments 1-4: one target.



(b) Experiments 5-6: two targets.



(c) Experiment 7: four targets.

Figure 2.4. Experiment arrangements, with respect to platform depth (pd) $d_0 = \frac{1}{4} pd$, $d_1 = \frac{1}{2} pd$, $d_2 = \frac{1}{6} pd$, $d_3 = \frac{1}{3} pd$, $d_4 = \frac{1}{4} pd$ and $d_5 = \frac{3}{4} pd$.

Chapter 3

Feature-sets

In this chapter, the features used throughout the thesis are reviewed. Compensation for target resolution and location effects can be implemented based on the concepts in [85], for simplicity we use the features without compensation.

The material presented in this chapter is partly taken from [87].

3.1 Motivation

Detection and classification for TWRI need to have reliable features that can be used at the pixel or voxel level. Using only a single polarization channel intensity as in [73, 75, 76], one can detect the targets to a certain level. In [84, 85] statistical and geometrical features are exploited to enhance the classification performance, but the approach used, make the classification result depend heavily on the first phase of target / non-target detection. This can lead to many false positives due to the small changes between the target pixels and its surroundings. Using multiple views [7, 72] or different polarization channel intensities [80] can enhance the target detection. As it is reasonable to assume that the more information or features are exploited, the better is the result, as long as sufficient training data is available so that the curse of dimensionality does not take any effect. Then providing more features about the pixels can enhance target detection and classification. Different polarimetric signatures are used in SAR applications [92–94] which implies they might be useful in TWRI as well. In SAR image analysis, many techniques have been developed that exploit spatial information [95–97]. The use of spatial models is mostly based on either textural features [98], contextual methods, like the Markov Random Field (MRF) model [99]. This motivates the use of features or methods that employ MRF neighborhood model, and can be used to extract texture features, like co-occurrence matrices [25, 99–101].

3.2 Polarization Signatures Sets

Polarimetric SAR images carry more information than the usual single polarized channel images [92–94]. Attempts to find an optimal combination of the information available from different polarimetric channels include [15–21].

Let \mathbf{X}_q , denote the vector of the complex polarimetric measurements at voxel site q . The full polarimetric vector for a reciprocal medium has three unique elements and is usually defined as [20] :

$$\mathbf{X}_q = \begin{pmatrix} S_{qHH} \\ S_{qHV} \\ S_{qVV} \end{pmatrix} \quad (3.1)$$

where S_{HV} is the complex amplitude of the H - (horizontal) polarized return given that the transmitted signal is V - (vertically) polarized. For a homogeneous region of a uniformly distributed scatterer, the fully polarimetric target properties can be described by the polarimetric covariance matrix [21]:

$$\mathbf{C}_q = E[X_q X_q^{-H}] = \begin{pmatrix} E(S_{qHH} S_{qHH}^*) & E(S_{qHH} S_{qHV}^*) & E(S_{qHH} S_{qVV}^*) \\ E(S_{qHV} S_{qHH}^*) & E(S_{qHV} S_{qHV}^*) & E(S_{qHV} S_{qVV}^*) \\ E(S_{qVV} S_{qHH}^*) & E(S_{qVV} S_{qHV}^*) & E(S_{qVV} S_{qVV}^*) \end{pmatrix} \quad (3.2)$$

where $E(\cdot)$ is the expected value and H defines the complex conjugate and transpose (Hermitian) operation. Hence, a feature vector that is derived from the parameters in the matrix \mathbf{C}_q can represent different classes on the imaged scenes [102]. For a practical implementation, the expected value $E(\cdot)$ is replaced by the sample mean. In this section, we will discuss some of the polarization signatures used in the literature and considered in our algorithms.

- (a) *Thresholded intensity features* : The automatic level thresholding algorithm [54] is used to threshold each complex component. The lower part is set to zero and the voxels with values above the threshold remain the same. The thresholded absolute value vector for all complex components is given as

$$\mathbf{T}_q = \begin{pmatrix} T_{qHH} \\ T_{qHV} \\ T_{qVV} \end{pmatrix} \quad (3.3)$$

- (b) *Normalized phase difference*: The phase difference between any two complex polarized components is computed as [93]:

$$\Psi_q = \arg(S_{qi} S_{qj}^*) \quad (3.4)$$

Usually the components considered are the two co-polarized components HH and VV [93].

- (c) *Normalized intensity ratio*: The intensity ratio between any two complex polarized components. [93]. We define the vector

$$\mathbf{IR}_{\mathbf{q}} = \begin{pmatrix} \frac{|S_{q\text{HH}}|^2}{|S_{q\text{VV}}|^2} \\ \frac{|S_{q\text{HH}}|^2}{|S_{q\text{HV}}|^2} \\ \frac{|S_{q\text{HV}}|^2}{|S_{q\text{VV}}|^2} \end{pmatrix} \quad (3.5)$$

for which the three features are referred to as $\text{IR}_{q\text{HHVV}}$, $\text{IR}_{q\text{HHHV}}$ and $\text{IR}_{q\text{HVVV}}$ respectively.

- (d) *Odd and even bounce*: The odd-bounce channel corresponds to the radar return from a flat plate or a trihedral; the even-bounce channel corresponds to the radar return from a dihedral. Few dihedral structures exist in natural clutter, but these structures are sufficient to describe most of the man-made targets [103–105]. Natural clutter tends to exhibit more odd bounce reflected energy,

$$E_{q(\text{odd})} = \frac{|S_{q\text{HH}} + S_{q\text{VV}}|^2}{2}, \quad (3.6)$$

than even-bounce reflected energy [103],

$$E_{q(\text{even})} = \frac{|S_{q\text{HH}} - S_{q\text{VV}}|^2}{2} + 2|S_{q\text{HV}}|^2. \quad (3.7)$$

- (e) *Real part and imaginary part of the normalized product*: In [21], many representations for the covariance matrix that characterize fully polarimetric data are introduced. These include features that mainly consist of intensities, phase differences, normalized or averaged products and the real and imaginary parts of the averaged product of different polarizations:

$$\mathbf{AP}_{\mathbf{q}(\text{RI})} = \begin{pmatrix} \text{Re}[E[S_{q\text{HH}}S_{q\text{VV}}^*]] \\ \text{Im}[E[S_{q\text{HH}}S_{q\text{VV}}^*]] \\ \text{Re}[E[S_{q\text{HH}}S_{q\text{HV}}^*]] \\ \text{Im}[E[S_{q\text{HH}}S_{q\text{HV}}^*]] \\ \text{Re}[E[S_{q\text{HV}}S_{q\text{VV}}^*]] \\ \text{Im}[E[S_{q\text{HV}}S_{q\text{VV}}^*]] \end{pmatrix} \quad (3.8)$$

We will refer to these features as $\text{Re}_{q\text{HHVV}^*}$, $\text{Im}_{q\text{HHVV}^*}$, $\text{Re}_{q\text{HHHV}^*}$, $\text{Im}_{q\text{HHHV}^*}$, $\text{Re}_{q\text{HVVV}^*}$ and $\text{Im}_{q\text{HVVV}^*}$ respectively.

- (f) *Logarithmic features*: in [102], the logarithm of the intensities, product magnitude and phase difference of different polarizations are suggested as character-

izing features:

$$\mathbf{IP}_{\mathbf{q}(\log)} = \begin{pmatrix} 10 \log_{10}(\mathbb{E}[|S_{qHH}|^2]) \\ 10 \log_{10}(\mathbb{E}[|S_{qVV}|^2]) \\ 10 \log_{10}(\mathbb{E}[|S_{qHV}|^2]) \\ 10 \log_{10}(|\mathbb{E}[S_{qHH}S_{qVV}^*]|) \\ 10 \frac{\arg(S_{qHH}S_{qVV}^*)}{\log(10)} \end{pmatrix} \quad (3.9)$$

The first three coefficients represent the backscatter cross sections of the surface element in decibel (dB) at three different linear polarizations. The fourth component measures the magnitude of the HH-VV correlation function, and the fifth one is proportional to the HH-VV phase difference. We will refer to these features as \log_{HH} , \log_{HV} , \log_{VV} , \log_{HHVV^*} , and \log_{PHHV^*} respectively. The advantages of operating in the log domain instead of the linear domain are two-fold. First, in the log domain, image speckle has the statistical characteristics of additive noise with a power level not varying much across the image, therefore, rendering clustering robust to the presence of image speckle. Second, the cross-polarized terms (i.e., HV) are often several orders of magnitude smaller than the co-polarized terms (i.e., HH or VV), which requires arbitrary weighting of the different channels. In the log-domain, the difference between the co-polarized and the cross-polarized terms is measured in dB and hence, is independent of the difference in absolute magnitude between each channel. Weighting of the different channels is thus not necessary [102].

- (g) *Spatial intensity features*: All the aforementioned features were suggested as features for TWRI on a previous work of the authors [87]. These features do not consider information about the pixel neighborhood. In this work, we introduce a new set of nonlinear features. This set exploits the output of a 3D median, maximum and minimum filters of the directly connected pixels neighborhood. This means the use of a $[3 \times 3 \times 3]$ window over each component of a polarization image.

$$\mathbf{SP}_q = \begin{pmatrix} \text{med}(|S_{qHH}|) \\ \text{max}(|S_{qHH}|) \\ \text{min}(|S_{qHH}|) \\ \text{med}(|S_{qHV}|) \\ \text{max}(|S_{qHV}|) \\ \text{min}(|S_{qHV}|) \\ \text{med}(|S_{qVV}|) \\ \text{max}(|S_{qVV}|) \\ \text{min}(|S_{qVV}|) \end{pmatrix} \quad (3.10)$$

This results in 9 features. We will refer to these features as med_{HH} , max_{HH} , and min_{HH} for features based on the HH polarization intensity. Features from

HV and VV polarizations will follow the same notation. We will refer to the vector of spatial intensity features as **SP**.

The proposed feature vector at each voxel q will contain the thresholded intensity features, logarithmic features, intensity ratio, real and imaginary parts of the averaged products, odd and even bounce returns and the spatial intensity features as follows:

$$\mathbf{FP}(q) = \begin{pmatrix} \mathbf{T}_q \\ \mathbf{IP}_{q(\log)} \\ \mathbf{IR}_q \\ \mathbf{AP}_{q\mathbf{RI}} \\ E_{q(\text{odd})} \\ E_{q(\text{even})} \\ \mathbf{SP}_q \end{pmatrix} \quad (3.11)$$

which contains 28 features in total. Table 3.1 list the features in brief.

3.3 Co-occurrence matrix

Due to the small variation between the target pixels and the closely surrounding noise and clutter pixels, targets become hard to detect. Algorithms that measure texture are built to deal with such small variations even in the case of non-optical images like SAR and MRI images [22]. Gray level co-occurrence matrix (GLCM), also known as co-occurrence distribution or gray-tone spatial dependence, is one of the well-known methods for texture features extraction [23–25]. The co-occurrence matrix contains texture, histogram and edge information [26] and was first proposed by Haralick *et al.* [106]. This matrix corresponds to the relative frequencies of gray level of pairs of pixels separated by a displacement vector $\Lambda = (\lambda_x, \lambda_y)$. With this representation the co-occurrence matrix is a $G \times G$ matrix where G is the total number of gray levels for a given image, and the gray level function for each pixel is $g(x, y)$. Thus, for an image I of size $N_x \times N_y$, the co-occurrence matrix can be generated as [107]:

$$C(i_1, i_2) = \frac{1}{N_x \times N_y} \sum_{x=0}^{N_x} \sum_{y=0}^{N_y} \begin{cases} 1, & \text{if } \{g(x, y) = i_1 \ \& \ g(x + \lambda_x, y + \lambda_y) = i_2\} \\ 0, & \text{otherwise} \end{cases} \quad (3.12)$$

Where i_1 and i_2 are gray levels that take values from 0 to $G - 1$. A sample co-occurrence matrix for a 4 x 6 image with 4 gray levels is shown in Figure 3.1.

Table 3.1. Polarization features, their corresponding position in the feature vector and the abbreviations to be used in the remainder of the thesis

Feature No.	Feature Definition and Symbol
1-3	Thresholded intensities T_{HH} , T_{HV} and T_{VV}
4-6	Logarithmic intensities \log_{HH} , \log_{HV} and \log_{VV}
7	Logarithmic absolute of the product $HH VV^* \rightarrow \log_{HHVV^*}$
8	Logarithmic phase of the product $HH VV^* \rightarrow \log_{PHHV^*}$
9-11	Intensity ratios IR_{HHVV} , IR_{HHHV} and IR_{VVHV}
12, 14, 16	Real part of the normalized products Re_{HHVV^*} , Re_{HHHV^*} and Re_{HVVV^*}
13, 15, 17	Imaginary part of the normalized products Im_{HHVV^*} , Im_{HHHV^*} and Im_{HVVV^*}
18-19	Odd and even bounce energies E_{odd} and E_{even}
20-28	Spatial intensity features med_{HH} , max_{HH} , min_{HH} , med_{HV} , max_{HV} , min_{HV} , med_{VV} , max_{VV} , and min_{VV}

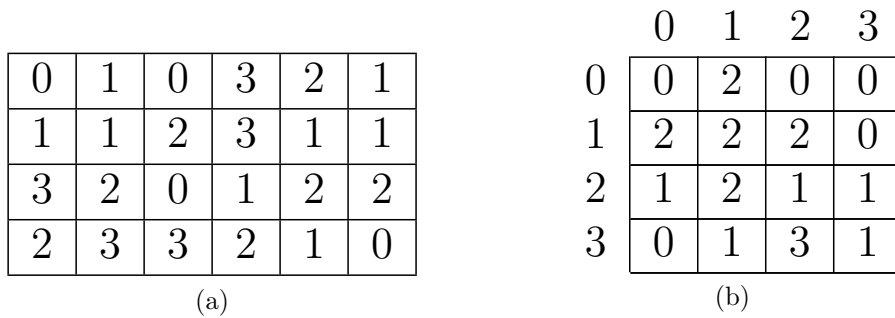


Figure 3.1. Sample co-occurrence matrix in (b) for a 4 x 6 image in (a) with 4 gray levels using a bivariate clique with distance vector $\Lambda = (1, 0)$.

In the literature, the 3D co-occurrence matrix or the 3D GLCM is used with different meanings. In [108], 3D refers to the image not to the co-occurrence matrix. In other works as in [101, 109], 3D refers to the co-occurrence matrix itself, meaning that the

gray level relative frequency of co-occurrence is calculated using three pixels instead of two as in the classical way. In our algorithm we extend the co-occurrence matrix to calculate the relative frequencies of gray levels between 2 to 5 pixels, meaning that we are dealing with two to five dimensional co-occurrence matrices. The relative frequencies are calculated for 3D images with different scanning arrangements of the pixels, or in this case, the voxels. We can consider the gray level of each voxel as a sample from a random variable, this leads to better representation of the joint probability of co-occurrence between neighboring voxels by the concept of Markov Random Fields (MRF). In [110] MRF is proposed as a distribution of conditional probabilities over elements in a lattice. This concept was introduced in [99,111–113] in the image processing context.

The voxels in a lattice \mathcal{L} of a 3D image are defined as a set of random variables $\mathcal{X} = \{X_1, X_2, \dots, X_R\}$, where each random variable X_r , $r = 1, 2, \dots, R$, is defined on a set $\mathcal{V} = \{0, 1, \dots, G-1\}$. A field is called MRF if it satisfies two conditions; the Markovianity, where the probability of an element having a certain value depends only on its neighbors $P(v_q|v_{\mathcal{L}-\{q\}}) = P(v_q|v_{\mathcal{N}_q})$ for a neighborhood system \mathcal{N} of an element q , \mathcal{N}_q , and the positivity where the joint probabilities of all realizations must have a positive value, $P(X) > 0 \forall X \in \mathcal{X}$. The joint probabilities are calculated over a subset of sites of the lattice \mathcal{L} called cliques. The cliques are governed by the neighborhood system used. The neighborhood system with the cliques that will be used throughout this thesis is the third-order Markov neighborhood [22] as shown in Figure 3.2. The considered neighborhood system has the configuration:

$$\mathcal{N}_q = \{q' \in \mathcal{L} | \text{dist}(v_q, v_{q'}) \leq 2, q' \neq q\} \quad (3.13)$$

where dist corresponds to the Euclidean distance.

The extension of the cliques shown in Figure 3.2 to 3D space are shown in Figure 3.3. The set $\Phi = \{\varphi_1, \varphi_2, \dots, \varphi_R\}$ include all the cliques used, where φ_r represents the clique φ with index r and R is the total number of cliques. In our set, there are altogether 54 cliques.

3.3.1 Multivariate Co-occurrence Matrices

The co-occurrence matrices can be considered as multivariate or k -variate p.d.f estimates (p.d.f.s), and according to the neighborhood system that we mentioned earlier, $k \in \{2, 3, 4, 5\}$ is the number of voxels in each clique. Each clique has a center or reference voxel, where the distance vector of all the other voxels in the arrangement

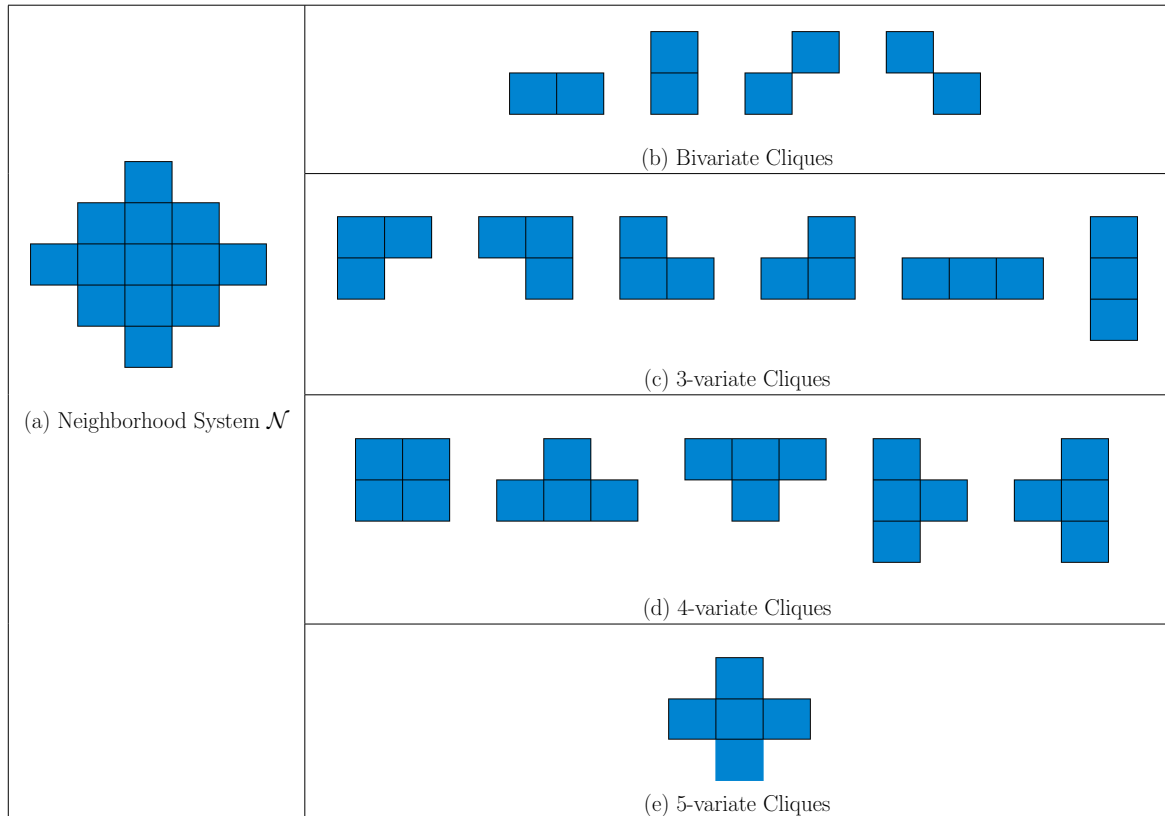


Figure 3.2. Directly connected MRF cliques for a third order neighborhood system. Figure redrawn from [22]

is calculated with reference to it. Let $g(x, y, z)$ denotes the gray level function of the reference voxel, then the extension of the co-occurrence matrix from bivariate to k -variate can be written as:

$$C_{\varphi_r}(i_1, \dots, i_k) = \frac{1}{\Xi} \sum_{\forall x, y, z} \begin{cases} 1, & \text{if } \{g(x, y, z) = i_1, \dots, g(x + \lambda_x^k, y + \lambda_y^k, z + \lambda_z^k) = i_k\} \\ 0, & \text{otherwise} \end{cases} \quad (3.14)$$

where Ξ is a normalizing factor such that $\sum_{\forall i_1, \dots, i_k} C(i_1, \dots, i_k) = 1$, i_k is the gray level of the voxel k on the clique φ_r , $\Lambda_k = (\lambda_x^k, \lambda_y^k, \lambda_z^k)$ is the distance of the voxel k from the reference voxel on the clique φ_r .

The k -variate co-occurrence matrices can be used to discriminate between different classes. For example, if we build separate co-occurrence matrices for target and clutter, we can use these matrices to classify between the two classes.

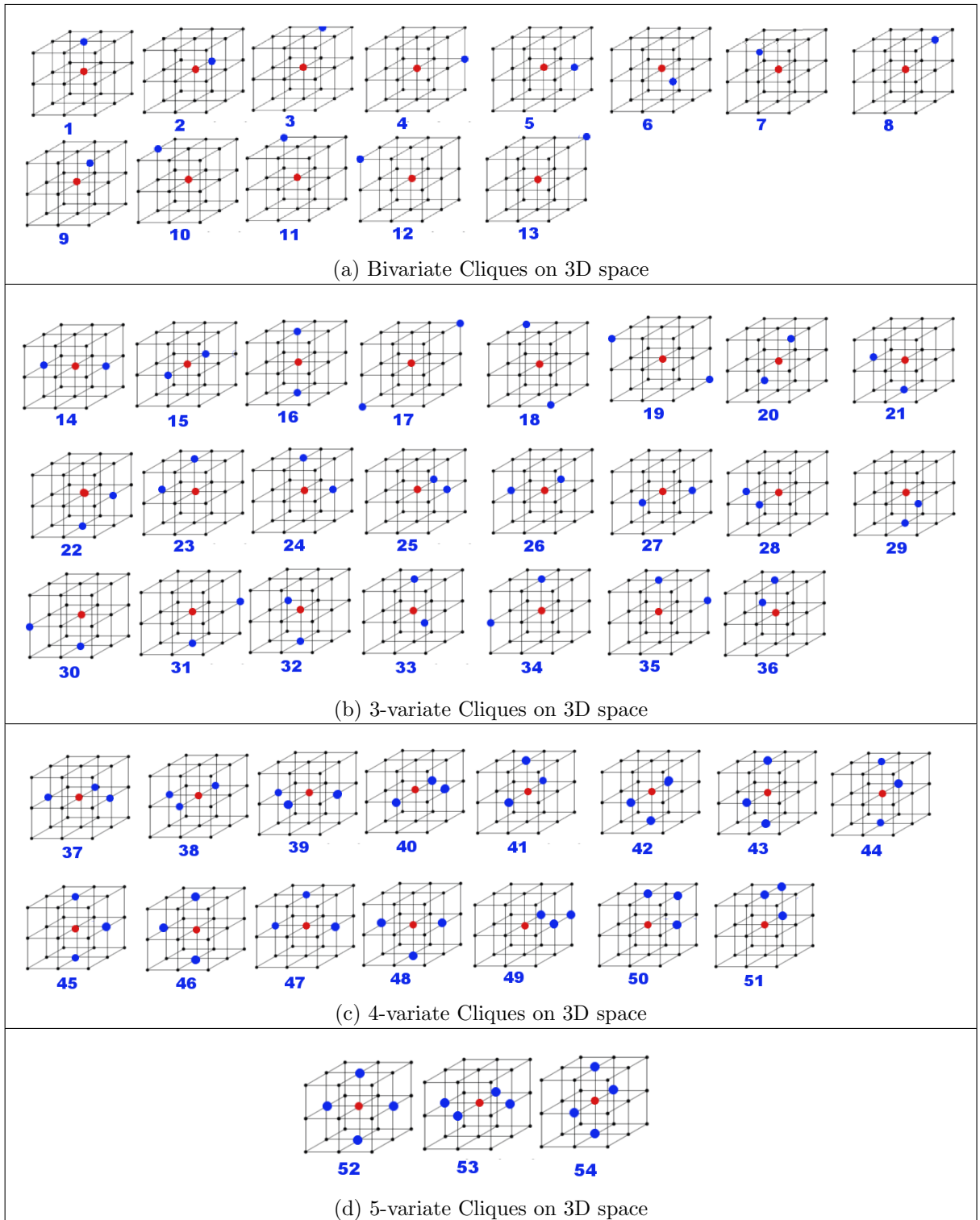


Figure 3.3. k -variate MRF cliques for a third order neighborhood system in 3D space.

3.3.2 Texture features from the multi-variate co-occurrence matrices

In [100,106,114], different texture features from the bivariate co-occurrence matrices were extracted. Here, we present modified equations that can be used to extract

some of these features from different k -variate co-occurrence matrices. The co-occurrence matrices here are not calculated one-time for the image as a whole, instead they are calculated through a moving window. Many variations for the size of the moving window can be implemented. In our case a 3D ($5 \times 5 \times 5$) window is used which is small enough not to be computationally expensive and large enough to allow the chosen cliques to move inside. At each transition of the moving window, the co-occurrence matrix is calculated for the voxels inside the window with all the different clique arrangements. For each k , we average all the calculated co-occurrence matrices, the resulting matrix is denoted by \bar{C} . The texture features are then extracted from \bar{C} and assigned to the feature vector for the image voxel at the center of the window. In this paper, we extended 9 conventional measures [106,115] from bivariate to multi-variate co-occurrence matrices. A novel formulation for extracting these features from k -variate co-occurrence matrices are shown in Table 3.2.

In the following we will use μ, σ^2 as the mean and variance of the normalized k -variate co-occurrence matrix, and μ_m, σ_m^2 as the mean and variance of the marginal co-occurrence matrix

$$C_m(i_m) = \sum_{i_1, \dots, i_k=0}^{G-1} C(i_1, \dots, i_k). \quad (3.15)$$

These features measure [115]:

1. Auto-correlation (*ACOR*): Measures the similarity between the gray level relative frequencies of the voxels.
2. Contrast (*CON*): Measures the differences between the highest and smallest gray level values of the set of voxels.
3. Correlation (*COR*): Measures the linear relationship between the gray levels of the voxels.
4. Dissimilarity (*DSM*): Measures the heterogeneity between gray level values of the voxels.
5. Angular second moment (*ASM*): Evaluates consistency of textural information.
6. Entropy (*ENT*): Measures the disorderliness or complexity of the texture.

Table 3.2. Extended equations for k-variate co-occurrence matrix texture features

Measure	for 2 variate co-occurrence matrix	for k -variate co-occurrence matrix
Auto-Correlation (<i>ACOR</i>)	$\sum_{i_1=0}^{G-1} \sum_{i_2=0}^{G-1} i_1 i_2 \overline{C}(i_1, i_2)$	$\sum_{i_1=0}^{G-1} \cdots \sum_{i_k=0}^{G-1} \prod_{u=1}^k i_u \overline{C}(i_1, \dots, i_k) \quad (3.16)$
Contrast (<i>CON</i>)	$\sum_{i_1=0}^{G-1} \sum_{i_2=0}^{G-1} (i_1 - i_2)^2 \overline{C}(i_1, i_2)$	$\sum_{i_1=0}^{G-1} \cdots \sum_{i_k=0}^{G-1} \overline{C}(i_1, \dots, i_k) \left[\sum_{\forall u \neq v} (i_u - i_v)^2 \right] \quad (3.17)$
Correlation (<i>COR</i>)	$\frac{\sum_{i_1=0}^{G-1} \sum_{i_2=0}^{G-1} (i_1 i_2) \overline{C}(i_1, i_2) - \mu_1 \mu_2}{\sigma_1 \sigma_2}$	$\frac{\sum_{i_1=0}^{G-1} \cdots \sum_{i_k=0}^{G-1} \prod_{u=1}^k i_u \overline{C}(i_1, \dots, i_k) - \prod_{u=1}^k \mu_u}{\prod_{u=1}^k \sigma_u} \quad (3.18)$
Dissimilarity (<i>DSM</i>)	$\sum_{i_1=0}^{G-1} \sum_{i_2=0}^{G-1} i_1 - i_2 \overline{C}(i_1, i_2)$	$\sum_{i_1=0}^{G-1} \cdots \sum_{i_k=0}^{G-1} \overline{C}(i_1, \dots, i_k) \left[\sum_{\forall u \neq v} i_u - i_v \right] \quad (3.19)$
Angular Second Moment (<i>ASM</i>)	$\sum_{i_1=0}^{G-1} \sum_{i_2=0}^{G-1} \overline{C}(i_1, i_2)^2$	$\sum_{i_1=0}^{G-1} \cdots \sum_{i_k=0}^{G-1} \overline{C}(i_1, \dots, i_k)^2 \quad (3.20)$
Entropy (<i>ENT</i>)	$\sum_{i_1=0}^{G-1} \sum_{i_2=0}^{G-1} \overline{C}(i_1, i_2) \log \overline{C}(i_1, i_2)$	$\sum_{i_1=0}^{G-1} \cdots \sum_{i_k=0}^{G-1} \overline{C}(i_1, \dots, i_k) \log \overline{C}(i_1, \dots, i_k) \quad (3.21)$
Homogeneity (<i>HOM</i>)	$\sum_{i_1=0}^{G-1} \sum_{i_2=0}^{G-1} \frac{\overline{C}(i_1, i_2)}{1 + (i_1 - i_2)^2}$	$\sum_{i_1=0}^{G-1} \cdots \sum_{i_k=0}^{G-1} \frac{\overline{C}(i_1, \dots, i_k)}{1 + \sum_{\forall u \neq v} (i_u - i_v)^2} \quad (3.22)$
Maximum Probability (<i>MP</i>)	$\max_{i_1, i_2} \overline{C}(i_1, i_2)$	$\max_{i_1, \dots, i_k} \overline{C}(i_1, \dots, i_k) \quad (3.23)$
Standard Deviation (<i>SD</i>)	$\sqrt{\sum_{i_1=0}^{G-1} \sum_{i_2=0}^{G-1} (i_1 \mu - \overline{C}(i_1, i_2))^2}$	$\sqrt{\sum_{i_1=0}^{G-1} \cdots \sum_{i_k=0}^{G-1} (i_1 \mu - \overline{C}(i_1, \dots, i_k))^2} \quad (3.24)$

7. Homogeneity (*HOM*): Measures the homogeneity between gray level values of the voxels.
8. Maximum probability (*MP*): Measures the maximum gray level relative frequency
9. Standard deviation (*SD*): Measures the degree of the gray level values distribution.

The vector of the texture features from the k -variate co-occurrence matrices contain ($9 \times 4 = 36$) features in total, which represent a set of nine features for each averaged k -variate co-occurrence matrix, where $k \in \{2, 3, 4, 5\}$. For $k = 2$ the feature vector, \mathbf{FC}_k , at voxel q is

$$\mathbf{FC}_2(q) = (ACOR_q \quad CON_q \quad COR_q \quad DSM_q \quad ASM_q \quad ENT_q \quad HOM_q \quad MP_q \quad SD_q)^T \quad (3.25)$$

and the whole vector of the texture features at voxel q is:

$$\mathbf{FC}(q) = \begin{pmatrix} \mathbf{FC}_2(q) \\ \mathbf{FC}_3(q) \\ \mathbf{FC}_4(q) \\ \mathbf{FC}_5(q) \end{pmatrix} \quad (3.26)$$

3.4 Conclusion

In this chapter, features for target detection and classification in TWRI are defined. Some polarimetric signatures that proved to provide effective representation on Synthetic Aperture Radar (SAR) applications are implemented. A New set of spatial polarimetric signatures is introduced. The estimated probabilities of co-occurrence matrices and co-occurrence matrices based texture features gave a good performance in remote sensing and radar applications. In this chapter, features based on co-occurrence matrices are extended to k -variate representation. The introduced feature-sets will be tested in the following chapters.

Chapter 4

Segmentation by classification

The contribution of this chapter is a framework for target detection using segmentation by classification, including exploitation of target polarization signatures. Each B-Scan of the 3D TWRI image is over-segmented to homogeneous regions without restriction on the shape of the region or the number of regions. These regions can be clustered together using agglomerative clustering to form three clusters corresponding to targets, clutter and noise classes. The resulting clustered or super-pixels data set is divided into a training set and a testing set. The training set is used to build a classifier that is based on the random forests algorithm [116], which is then tested using the testing data set. A diagram summarizing the general steps of the proposed algorithm is shown in Figure 4.1.

The material presented in this chapter is partly taken from [87].

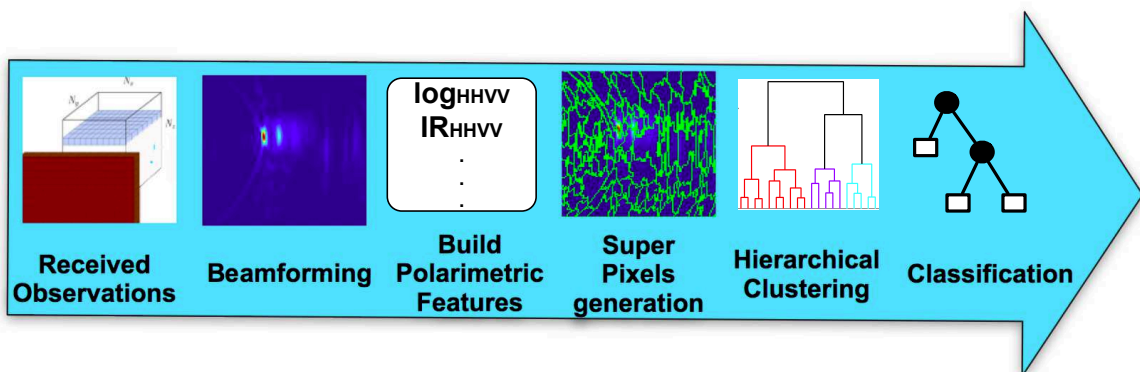


Figure 4.1. The general steps of Segmentation by classification.

4.1 Motivation

All existing image-domain target detection algorithms in TWRI use a pixel-grid as the underlying representation, cf. [7, 73, 75, 76, 84]. However, the pixel-grid is not a natural representation of visual scenes. It is rather an “artifact” of a digital imaging process [117, 118]. It would be more natural, and more efficient, to work

with perceptually meaningful entities, obtained from a low-level grouping process [117, 119]. For that, we apply over-segmentation to segment the image into regions with homogeneous characteristics. Subsequently, for classification of super-pixels, decision trees are considered. Decision trees are one of the most popular classification algorithms used in data mining and machine learning to create knowledge structures that guide the decision making process [120–122]. Thus, they have the ability to handle high-dimensional data well and to ignore irrelevant features [123]. These are desirable properties for the problem at hand.

4.2 Segmentation

In this section, we present a pre-processing stage to group pixels into homogeneous regions called ‘super-pixels’. To create the super-pixels for our TWRI B-Scans, we used the quick shift algorithm [124] for segmentation. Unlike other super-pixelization schemes, such as the ones based on normalized cuts [117], quick shift produces super-pixels that are not fixed in approximate size or number. A complex image with many fine scale image structures may have many more super-pixels than a simple one, and there is no parameter which puts a penalty on the boundary, leading to super-pixels which are quite varied in size and shape [118].

From Equation (2.2) of Section 2.1, we can get the complex B-Scan image I for each height. Combining the different B-Scans for all heights will give a 3D image for the whole scene. Each pixel has an associated vector of features as described in Section 3.2. The feature vector includes all the proposed polarimetric signatures except the spatial ones. The feature vector associated with each pixel as in Equation (3.11) can be identified as $\mathbf{F}_q(d)$ where $d = 0, \dots, D - 1$ and D is the number of features. Quick shift is a mode seeking algorithm which links each pixel to its nearest neighbor which has an increase in the estimate of the density. These links form a tree where the root of the tree is the pixel, which corresponds to the highest mode in the image [124]. Figure 4.2 shows a simple illustration of how quick shift forms the tree. Quick shift regards each pixel q , $q = 0, \dots, Q - 1$, as a sample from a $D + 2$ dimensional vector space. It then calculates the Parzen density estimate (with a Gaussian kernel of standard deviation σ) [125]:

$$E(q) = P(x_q, y_q, \mathbf{F}_q) = \sum_{j=0}^{J-1} \frac{1}{(2\pi\sigma)^{D+2}} \exp \left(-\frac{1}{2\sigma^2} \begin{bmatrix} x_q - x_j \\ y_q - y_j \\ \mathbf{F}_q - \mathbf{F}_j \end{bmatrix}^T \begin{bmatrix} x_q - x_j \\ y_q - y_j \\ \mathbf{F}_q - \mathbf{F}_j \end{bmatrix} \right) \quad (4.1)$$

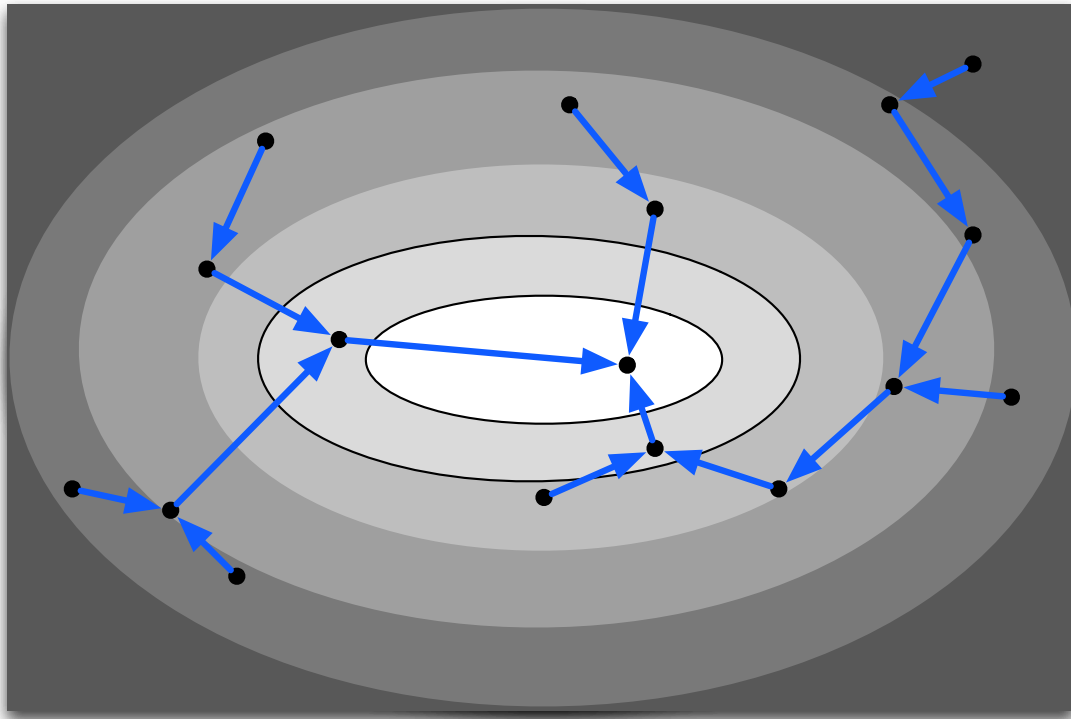


Figure 4.2. Tree forming illustration using quick shift. The black dots represent (some of) the data points, and the intensity of the image is proportional to the Parzen density estimate $E(q)$ from equation (4.1). Redrawn from [124]

where $j = 0, 1, \dots, J - 1$ denotes the neighboring pixel j , J is the number of neighboring pixels, x and y are the physical coordinates of the pixel. Then quick shift constructs a tree of pixels, connecting each image pixel to its nearest neighbor which has a greater density value. Formally, $E(j) > E(q)$ if, and only if [124, 125]

$$P(x_j, y_j, \mathbf{F}_j) > P(x_q, y_q, \mathbf{F}_q). \quad (4.2)$$

Each pixel (x_q, y_q) is connected to the closest higher density pixel parent (x'_q, y'_q) that achieves the minimum distance in:

$$\text{dist}(q) = \min_{P(j) > P(q)} ((x_q - x_j)^2 + (y_q - y_j)^2 + \|\mathbf{F}_q - \mathbf{F}_j\|_2^2) \quad (4.3)$$

The algorithm connects all the points into a single tree. Modes are then recovered by breaking the branches of the tree that are longer than a threshold τ . We used the quick shift implementation in the library VLFeat [125]. A typical segmentation result using quick shift applied to B-Scans of one and two target scenes using the feature set described in Section 3.2 is shown in Figure 4.3.

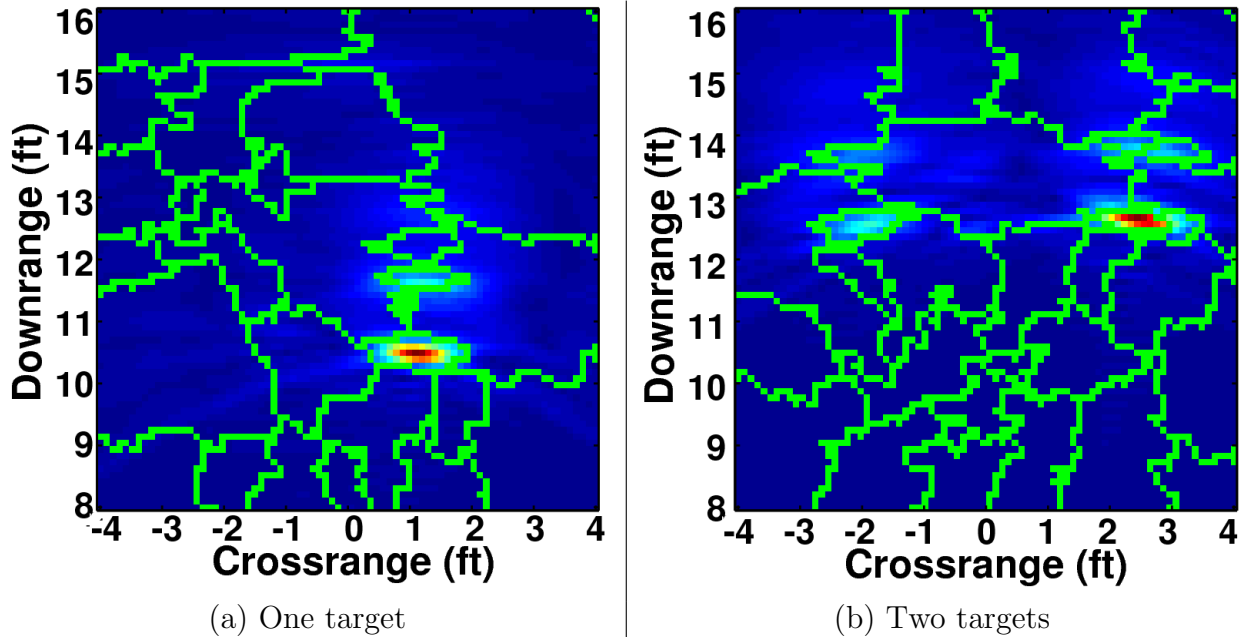


Figure 4.3. Typical results for quick shift segmentation of the images from Figure 2.2

4.3 Clustering

The aim is to cluster the super-pixels resulting from the previous stage to a maximum of three clusters for each B-Scan, corresponding to targets, clutter and noise. We note that clustering is an optional step and may not be used if the number of super-pixels is low. In this case, the super-pixels can be used directly for training and testing of the classifier presented in the next section. To achieve clustering, we first calculate a single feature vector for each super-pixel based on the features of all underlying pixels contained in a super-pixel. We take the median for each feature throughout all pixels that belong to this super-pixel to be the new value of this feature for this super-pixel. If ξ_s is the set of pixels belonging to super pixel s , then the new value for the feature d in the super-pixel feature vector \mathbf{FS}_s is:

$$\mathbf{FS}_s(d) = \tilde{\mathbf{F}}_{q \in \xi_s}(d) \quad (4.4)$$

where s is the super-pixel index, $s = 0, \dots, S - 1$, S is the number of super-pixels, $q \in \xi_s$ denotes the group of pixels that belong to super-pixel s , $\tilde{\mathbf{F}}(d)$ is the median of feature d over this group of pixels, and $d = 0, 1, \dots, D - 1$ is the feature index.

We use hierarchical agglomerative clustering [126, 127] to group the super-pixels into clusters, using the Euclidean distance to calculate the distances between the

super-pixels, i.e. the distance between two feature vectors \mathbf{FS}_1 and \mathbf{FS}_2 is given by

$$\|\mathbf{FS}_1 - \mathbf{FS}_2\|_2 = \sqrt{\sum_{d=0}^D (\mathbf{FS}_1(d) - \mathbf{FS}_2(d))^2}. \quad (4.5)$$

Subsequently, a linkage function is used to group the super-pixels hierarchically into clusters, a single linkage [126, 127] (also known as nearest neighbor or shortest distance) method is used to calculate the distance matrix between elements. The result of linkage can be represented visually as a dendrogram. The numbers along the horizontal axis represent the indices of the super-pixels. The upside-down U-shaped lines represent the links between super-pixels. The height of U indicates the distance between the super-pixels. A typical dendrogram result from clustering the two target image from Figure 4.3 is shown in Figure 4.4.

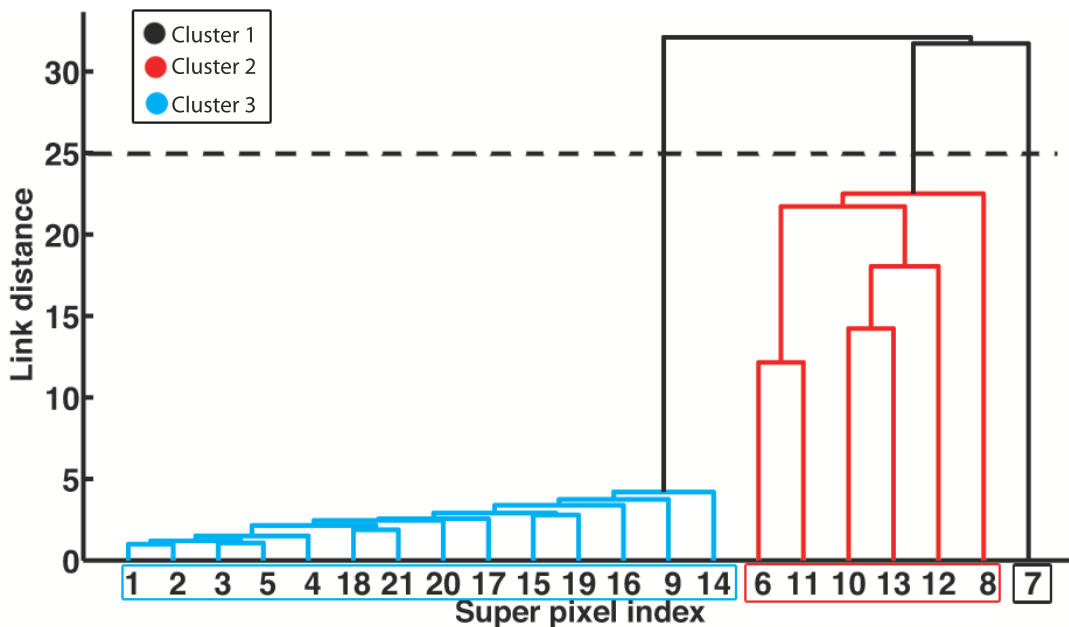


Figure 4.4. A typical dendrogram to visualize clustering for a two target B-Scan. Different colors represent different clusters. The horizontal dashed line illustrates where to cut to get 3 clusters

4.4 Classification with Random Forests

In this section, we propose a sequential classification scheme using random forests ensemble trees classifier [116]. Previous work regarding stationary target classification for TWRI includes [62, 86] where a minimum Mahalanobis distance classifier is

used to classify targets using 3D statistical models. Further, in [84] Iterative Conditional Modes (ICM) segmentation was considered to divide the TWRI image into target and noise, then nearest-neighbor classifier and support vector machines were used for further classification using superquadrics features.

In our classification problem, we are given data on a set of N super-pixels for training, $\chi = \{(\mathbf{FS}_1, O_1), \dots, (\mathbf{FS}_N, O_N)\}$, where \mathbf{FS}_n is a vector of descriptors and O_n is the corresponding n -th class label. Our goal is to find a model for predicting the values of O from new \mathbf{FS} values. Classification tree methods yield rectangular sets A_j such that the predicted value of O is j if \mathbf{FS} belongs to A_j for $j = 1, 2, \dots, J$ where J is the number of disjoint sets. By recursively partitioning the data set one variable from \mathbf{FS} at a time, the splits resulting from using each of the predictors are examined. The best split that maximizes homogeneity for both its parts is chosen. Then this procedure is repeated until there is no more possible splits or a stopping criterion is met. Deciding for the ‘best’ split is still an active research area, the most used method is the Classification And Regression Trees (CART) [120] which measures the impurities at each split using the Gini [128] impurity algorithm. Let t_p be a parent node, t_l, t_r the left and right child nodes respectively, χ a data set with D descriptors, N the number of observations, and K the number of classes. At each split, we have \mathbf{FS}_d^R which is the best threshold for the descriptor \mathbf{FS}_d that gives maximum homogeneity for each part. At the parent node t_p , the impurity is constant for all possible splits, so the change of impurity for the left and right nodes is calculated as follows [129]:

$$\Delta i(t) = i(t_p) - P_l(t_l) - P_r(t_r) \quad (4.6)$$

where P_l and P_r are the probabilities of the left and the right nodes. The objective is to solve for

$$\arg \max_{\mathbf{FS}_d \leq \mathbf{FS}_d^R} \Delta[i(t)] = \arg \max_{\mathbf{FS}_d \leq \mathbf{FS}_d^R} [i(t_p) - P_l(t_l) - P_r(t_r)] \quad (4.7)$$

This means that all possible values of the descriptors will be searched for the best split $\mathbf{FS}_d \leq \mathbf{FS}_d^R$ that will maximize the change of impurity $\Delta i(t)$. As mentioned before, the Gini function [128] will be used as impurity function here, which is given by:

$$i(t) = 1 - \sum_{k=1}^K p^2(k|t) \quad (4.8)$$

where $k = 1, 2, \dots, K$ is the class index and $p(k|t)$ is the probability of class k given node t . Substituting Equation (4.8) in Equation (4.7) gives:

$$\arg \max_{\mathbf{FS}_d \leq \mathbf{FS}_d^R} \Delta[i(t)] = \arg \max_{\mathbf{FS}_d \leq \mathbf{FS}_d^R} \left[- \sum_{k=1}^K p^2(k|t_p) + P_l \sum_{k=1}^K p^2(k|t_l) + P_r \sum_{k=1}^K p^2(k|t_r) \right] \quad (4.9)$$

which is the problem to be solved at each split throughout all possible values of tested descriptors. A block diagram illustrating the process is shown in Figure 4.5. A typical decision tree built on a database of around 200 B-Scans for scenes that contain one target (different types of targets) is shown in Figure 4.6. The major drawback, however, is that decision trees usually have relatively low prediction accuracy [123]. One of the best ways to improve the performance of Decision Tree-based algorithms is to use ensembles of trees [130] and Random Forests are one of the methods to enhance performance [116, 123].

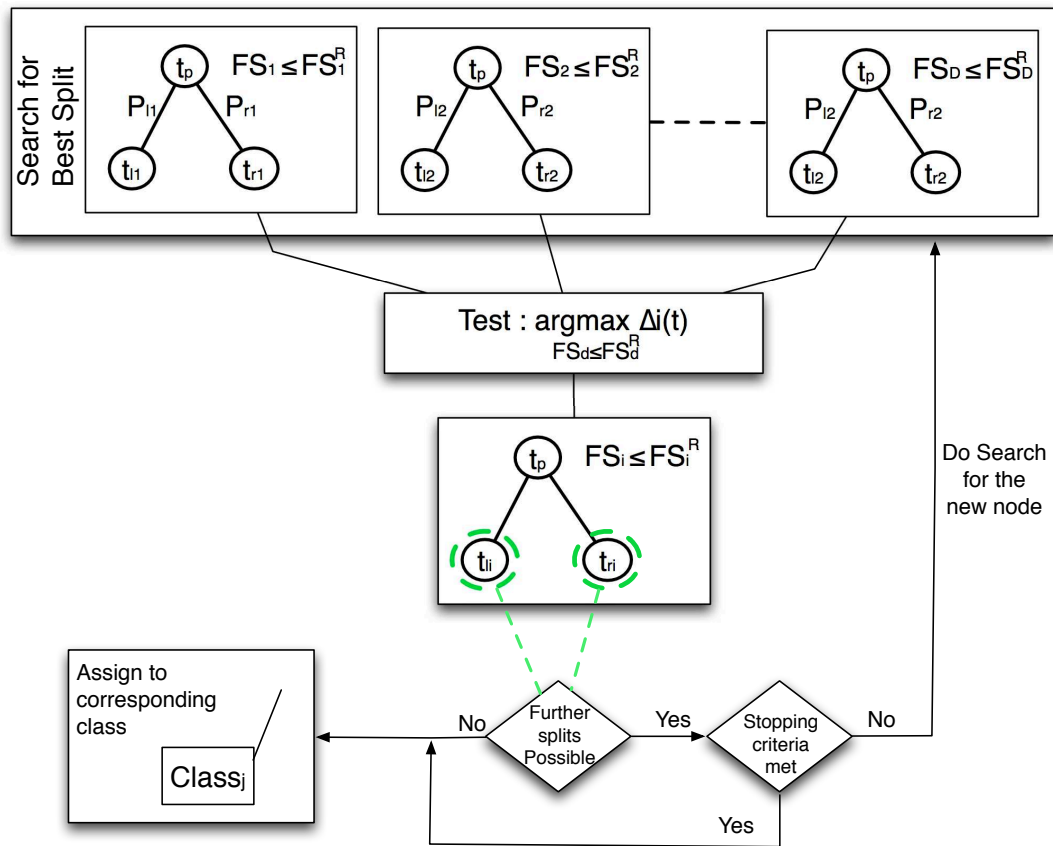


Figure 4.5. Decision tree growing process

A Random Forest is an ensemble of trees $T_0(\mathbf{FS}), \dots, T_{B-1}(\mathbf{FS})$, where \mathbf{FS} is the D -dimensional vector of features or descriptors associated with a super-pixel (or with a cluster if clustering is used) and B is the number of trees in the ensemble. The ensemble produces B outputs $\hat{O}_0, \hat{O}_1, \dots, \hat{O}_{B-1}$ where \hat{O}_b is the prediction by the b th tree and $b = 0, 1, \dots, B - 1$. Outputs of all trees are aggregated to produce one final prediction \hat{O} , which is the class predicted by the majority of trees. Given data on a set of N super-pixels for training, $\chi = (\mathbf{FS}_0, O_0), \dots, (\mathbf{FS}_{N-1}, O_{N-1})$, where \mathbf{FS}_n is

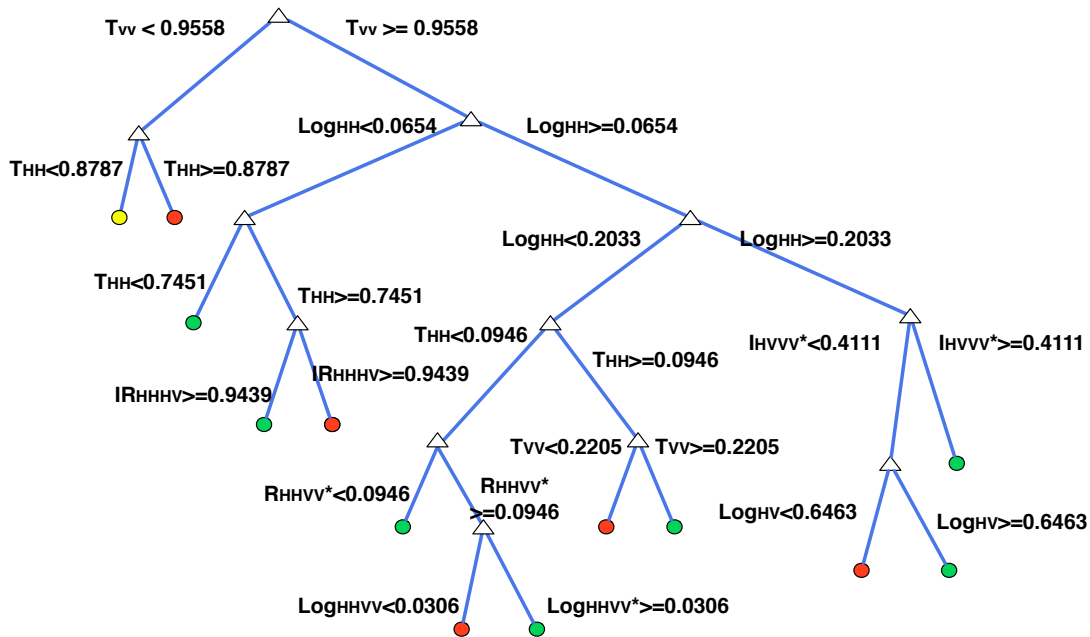


Figure 4.6. A typical Decision tree for one-target scene, yellow dots for the noise class, red is for the clutter class, and green is for the target class.

a vector of descriptors and O_n is the corresponding class label. Then, the following steps are implemented for training:

1. From N super-pixels as training data, randomly sample, with replacement (bootstrap [131]), to create B bootstrap samples (same number of trees).
2. For each bootstrap sample, grow a decision tree using only α randomly selected features to test for best splitting at each node (rather than all features). The tree is grown to the maximum size (until no further splits are possible) and not pruned back.

A typical block diagram illustrating the training and testing process using Random Forest(s) is shown in Figure 4.7. Random Forests perform a type of cross-validation in parallel with the training step by using Out-Of-Bag (OOB) samples [132]. Specifically, in the process of training, each tree is grown using a bootstrap sample. Since bootstrapping is sampling with replacement from the training data, some of the super-pixels will not be included while others will be repeated on each sample. The super-pixel samples that were not included are called 'Out-Of-Bag' samples. Usually each tree is grown using two-thirds of the training data and one-third is the OOB

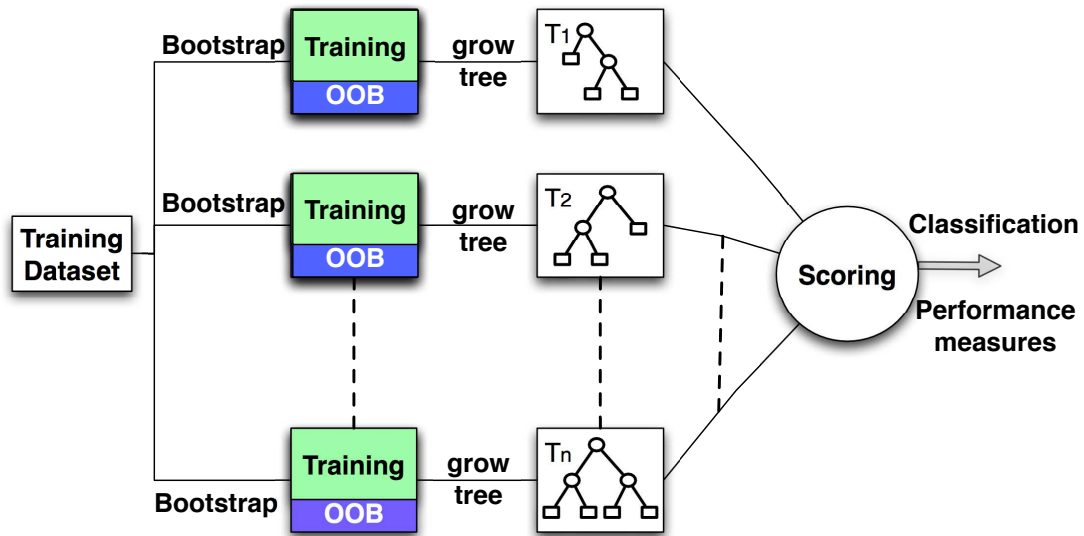
samples. These OOB samples can be used to estimate the ensemble prediction performance by testing the b^{th} tree T_b against χ_b^{OOB} and calculating the classification error as follows:

$$ER \approx ER^{OOB} = \frac{\sum_{i=1}^l I(\hat{O}^{OOB}(\mathbf{FS}_i) \neq O_i)}{l} \quad (4.10)$$

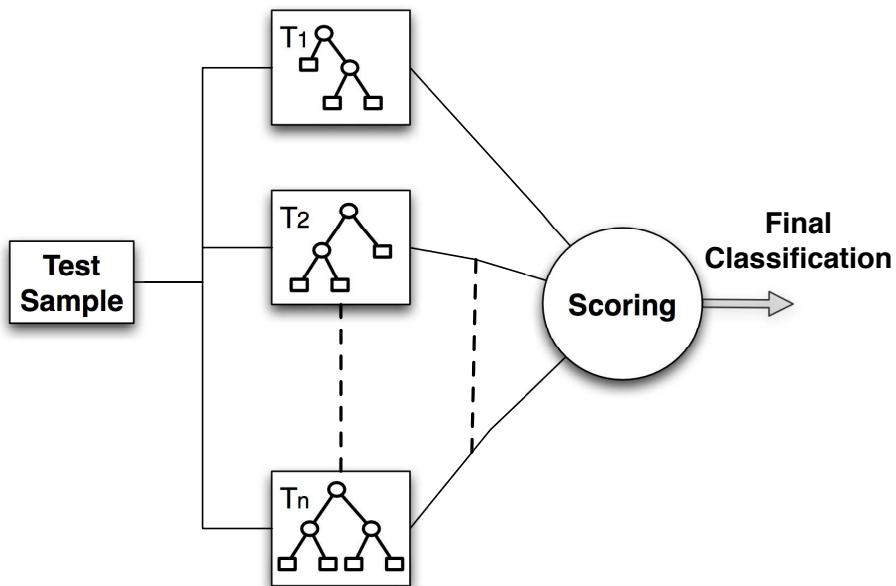
where $I(\cdot)$ is an indicator function, and l is the number of samples used [123]. A measure of how each feature contributes to the prediction accuracy can be calculated in the training process too. When a feature is 'noised up' (e.g., replaced with random noise), the accuracy of prediction should noticeably degrade if the feature contributes to the prediction accuracy. On the other hand, it should have little effect on the performance if it is irrelevant. In classification, the change in prediction accuracy is usually a less sensitive measure than the change in the margin. The margin is the difference between the proportion of votes for the correct class and the maximum of the proportion of votes for the incorrect classes. The tendency of the margin to become smaller (more negative or less positive) when a feature is 'noised up' is what is used to assess the feature importance in classification [123].

The feature importance is calculated as follows: use the OOB data part χ_b^{OOB} to make predictions of the b^{th} tree when this tree is grown. At the same time, each feature in the χ_b^{OOB} data is randomly permuted, one at a time, and predicted by the tree. At the end of the training process, the margins for each class are calculated based on both the original OOB and the OOB data with each feature permuted. Then the measure of importance for the d^{th} feature is simply $\beta - \beta_d$ where β is the average margin based on the OOB prediction and β_d is the average margin based on the OOB prediction with the d^{th} feature permuted.

After calculating the features importance, a new ensemble of trees is built upon the most contributing features only. This can be used to build different ensembles according to the classes needed, for example, an ensemble can be built to discriminate between target and non-target super-pixels. Another one can be built to discriminate between metal and non-metal targets, or between 'dihedral, sphere, and trihedral' classes. This may require additional types of features like shape features for example, but here, we restrict ourselves to polarimetric features only. These ensembles could be used sequentially or in parallel or have mixed architectures. A proposed scheme for the sequential classifier is shown in Figure 4.8. The decisions for each classifier could be semantically grouped to provide a better understanding about the target. Furthermore, a third level could be added on the scheme provided on Figure 4.8 to deal with the 3D image directly, which could make use of geometrical and shape features to add more information about the targets detected.



(a) Training Process



(b) Testing process

Figure 4.7. Random forests training and testing process.

4.5 Experimental Results

4.5.1 Experiments description

We consider the scenario presented in Section 2.1 for the evaluation of the proposed techniques. As mentioned we have four ‘one-target’ 3D images and two ‘two-target’

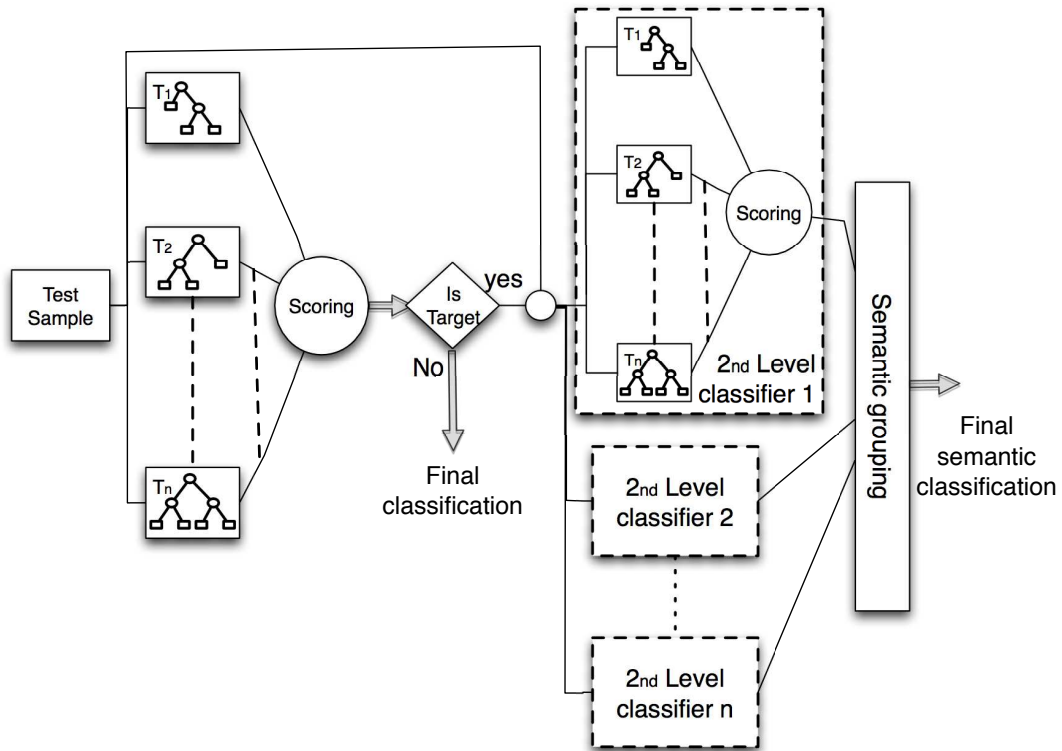


Figure 4.8. Block diagram for the multi-level classifier.

3D images, each 3D image is constructed from a set of 57×57 2D B-Scans. Each of these images has 3 different polarization data images, 'HH, HV and VV' resulting in 18 different sets. Our experiments are arranged as follows (a summary is given in Table 4.1)

- (a) Prepare the polarimetric data of the images.
- (b) Get the polarimetric features or signatures for the images. Each pixel will have a vector of 19 features as mentioned in Section 3.2.
- (c) The feature matrix should be checked for outliers and scaled so as to avoid unexpected behavior.
- (d) Use Quickshift to segment the images based on the polarimetric features. Each image will be segmented to regions called super-pixels as shown in Figure 4.3.
- (e) Clean the resulting super-pixels by removing super-pixels that contain only one pixel.

- (f) For each super-pixel, get the median of the feature vectors of all the super-pixels contained on this super-pixel. This will be the new feature vector for this super-pixel as mentioned in Section 4.2.
- (g) Optionally use clustering if the number of the super-pixels is large or the number of super-pixels that contain noise and clutter classes is much larger than super-pixels that contain target classes as mentioned in Section 4.3. If clustering is used, then again a median feature vector should be generated for each cluster based on the super-pixels contained on this cluster.
- (h) Separate the data into training and testing sets. The training set is two-thirds, and the testing set is one-third of the data.
- (i) Prepare three training data groups. The first uses the ‘one-target’ data, the second the ‘two-target’ data and the third is a mixed group between one and two target data. Also two testing data groups based on ‘one-target’ and ‘two-target’ data are prepared.
- (j) Use the training data to train the first ensemble of trees to classify between noise, clutter, and target classes. Subsequently re-build the ensemble using the most effective features only.
- (k) For the three groups of the training data, the previous step should be repeated to create three different ensembles.
- (l) The last steps are repeated to build two other groups of ensembles. The first classifies between generic ‘metal’ and ‘non-metal’ classes and the second classifies between the target types ‘dihedral’, ‘sphere’, ‘trihedral’, and ‘salt-water jug’.

4.5.2 Target, noise and clutter classification

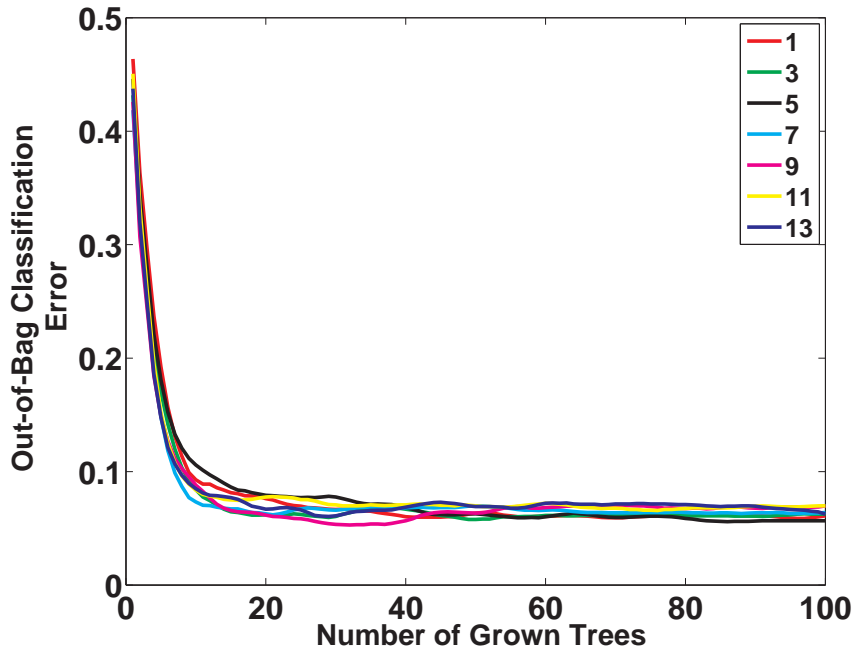
The first ensemble of 100 trees is trained to classify between three classes (‘target’, ‘noise’, and ‘clutter’) with all the features used. As mentioned before, three different ensembles depending on the training data (‘one-target’, ‘two-target’, and ‘mixed’) are grown. We first need to choose the number of features that will be randomly tested to find the best split at each tree node. We conducted different experiments to choose the optimum number of features tested to find the best split using the numbers = (1, 3, 5, 7, 9, 11, 13) as candidates. Both the classification error and the mean classification margin against the number of grown trees have been checked. The number of features that gave the lowest OOB classification error and best mean

Table 4.1. Segmentation by classification algorithm steps.

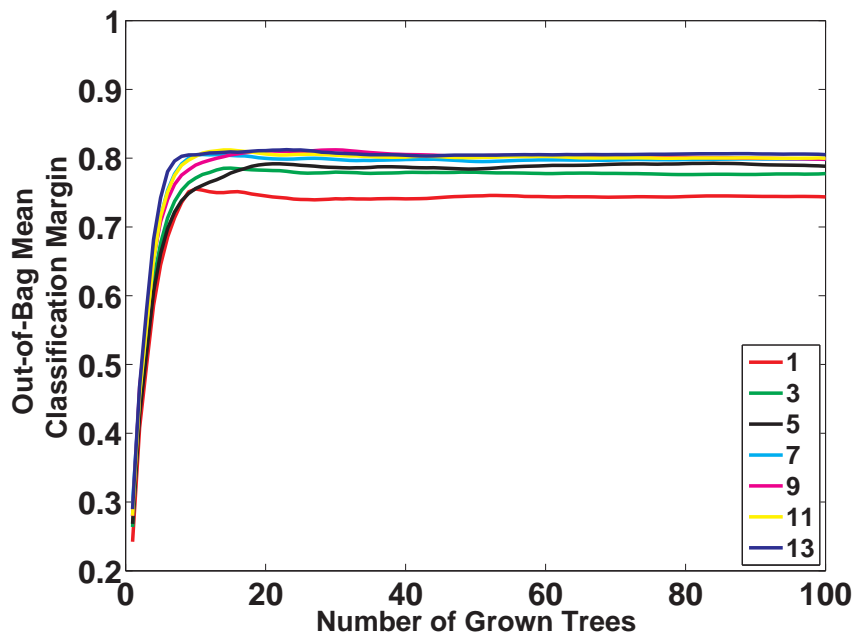
No.	Step
1	Prepare polarimetric images. A set of B-scans for each polarization. Each B-Scan has Q pixels.
2	Extract feature vector F_q with D features for each pixel q .
5	Generate S super-pixels for each B-Scan using quick shift segmentation.
4	Generate Feature vector FS_s for each super-pixel s using the underlying pixels features.
5	Prepare training and testing data for classification.
6	Do preparatory training for Random forests to choose best number of features for splitting and most important features.
7	Build RF classifier with these parameters.
8	Repeat for different purpose classifiers (target or not, human or not, metal or not, ..etc) and arrange in multi-level scheme.

classification margin is used. Figure 4.9 shows the OOB classification error and the mean classification margin for the three ensembles. The lowest classification error is gained using 3, 5, 7, and 13 random features at each split on the different ensembles, and among these numbers, the number of features that resulted in the highest mean classification margin is chosen for each case. Thus, 7 features for the case of ‘one-target’ data and 5 features for the other two cases are chosen for testing for best split. To estimate the feature importance, the features are permuted and tested as mentioned in Section 4.4. The results are shown in Figure 4.10. As can be seen, the six most important features for the ensemble trained with ‘one-target’ data are (1, 3, 4, 6, 10, 15), which are: the thresholded intensity for HH (T_{HH}), the thresholded intensity for VV (T_{VV}), the logarithmic intensity for HH (\log_{HH}), the logarithmic intensity for VV (\log_{VV}), and the imaginary part of the product of HH and HV* (Im_{HHHV^*}). The features with importance above 0.6 are (1, 4, 6). A comparison between the OOB classification error for both the reduced feature sets and the set with all features is shown in Figure 4.11, which shows no significant degradation in performance for both of the reduced feature sets. Thus, a feature set of the 6 most

important features will be used. Table 4.8 lists the most important features for each case.

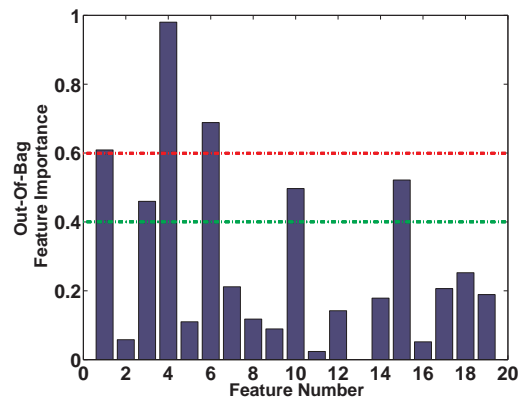


(a)

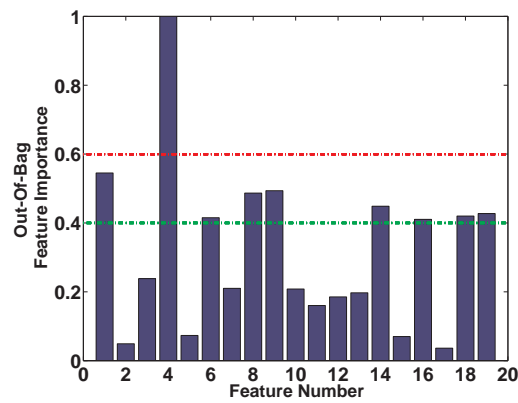


(b)

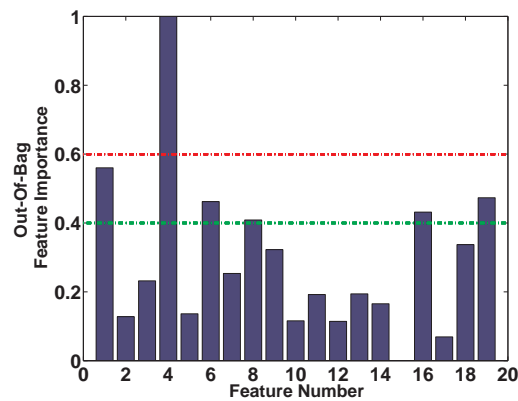
Figure 4.9. Test for the suitable number of random features at each split. Top is the OOB classification error, and the bottom is the OOB classification margin, both versus the number of grown trees. Ensembles are trained on mixed data.



(a)



(b)



(c)

Figure 4.10. Test for the feature importance for classifying ‘target’, ‘clutter’ and ‘noise’ classes. The green and red line are guide-lines reflect a feature importance of 0.4 and 0.6 respectively. The plot in (a) is for ensemble trained on ‘one-target’ data, (b) is for ensemble trained on ‘two-target’ data, and (c) is for ensemble trained on mixed data.

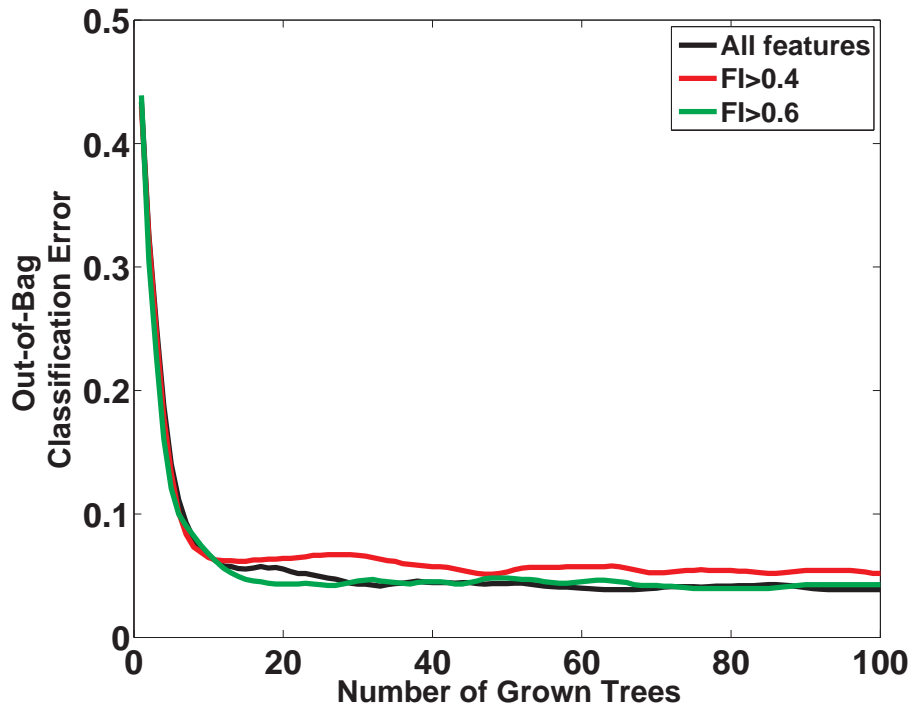


Figure 4.11. OOB Classification error for reduced feature set compared to using all features

Receiver operating characteristics (ROC) [133] performance curves for the reduced feature set of the ensemble trained on ‘one-target’ data are shown in Figure 4.12. The left curve depicts the false positive rate against the true positive rate. The value of the Area Under the Curve (AUC) is 0.9887 and an optimal operating point at (0.0486, 0.9753). The right curve is the ensemble accuracy vs. threshold on the score for the ‘target’ class. The curve shows a flat region indicating that any threshold from 0.3 to 0.6 is a reasonable choice. Calculating the maximum accuracy shows that it is achieved at a threshold around 0.45. Table 4.2 lists the AUC, the optimal operating point and the optimal threshold for all cases.

Testing the different ensembles of trees, which was trained using samples from the ‘one-target’, ‘two-target’ and ‘mixed’ data with both the ‘one-target’ data and the ‘two-target’ data gives a classification error between 0.04 – 0.072 for matched cases, around 0.1 for the mismatched case, and between 0.03 and 0.043 for the mixed training case. The classification error for all cases is listed in Table 4.3 in comparison to the Mahalanobis distance classifier (using Iterated conditional modes (ICM) and Levelset (LSM) segmentation methods) [84].

Figure 4.13 demonstrates the distance between the ‘target’, ‘noise’ and ‘clutter’

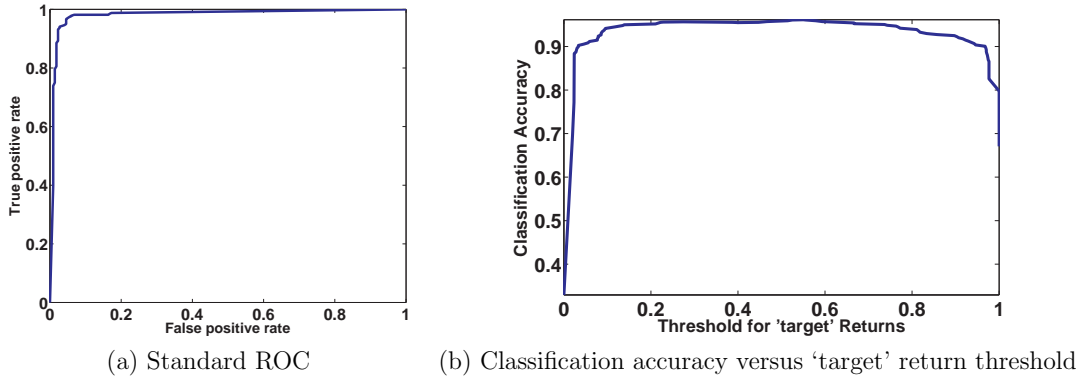


Figure 4.12. ROC curve and classification accuracy for a reduced feature set

Table 4.2. Performance measures for the first ensemble using OOB samples, classifying ‘target’, ‘noise’, and ‘clutter’ classes

Training \ Measure	AUC	Optimal threshold	Optimal operating point
one target	0.9887	0.45	(0.0486, 0.9753)
two targets	0.9896	0.5279	(0.0280, 0.9247)
mixed	0.9878	0.406	(0.0668, 0.9685)

classes using the most affecting features for the ensemble trained with mixed data. The three lines for each class represent the median and quantiles at 0.1 and 0.9 of the observations. It is clear that the first three features discriminate well between the different classes and the other features could help tuning after the different tree splits.

4.5.3 Further classification, metal/non-metal and target type classification

Repeating the previous experiments changing the classification objective as mentioned before. First, discriminate between the classes ‘metal’ and ‘non-metal’. Then, between the classes ‘dihedral’, ‘sphere’, ‘trihedral’, and ‘salt water jug’. Experiments show that the proposed scheme can be used for such a classification task. Figure 4.14 shows the feature importance when classifying for metal targets. Table 4.4 lists the AUC, the optimal operating point and the optimal threshold for this case.

Table 4.3. Classification errors for the first ensemble, classifying ‘target’, ‘noise’, and ‘clutter’ classes. Segmentation by classification, and Mahalanobis distance classifier (using ICM and LSM segmentation methods) are compared.

		Seg. by Class.			ICM			LSM		
		one target	two target	mixed	one target	two target	mixed	one target	two target	mixed
Testing	one target	0.043	0.116	0.0429	0.063	0.176	0.0735	0.0763	0.2263	0.1093
	two targets	0.139	0.072	0.0389	0.343	0.0763	0.0459	0.3433	0.2216	0.0863

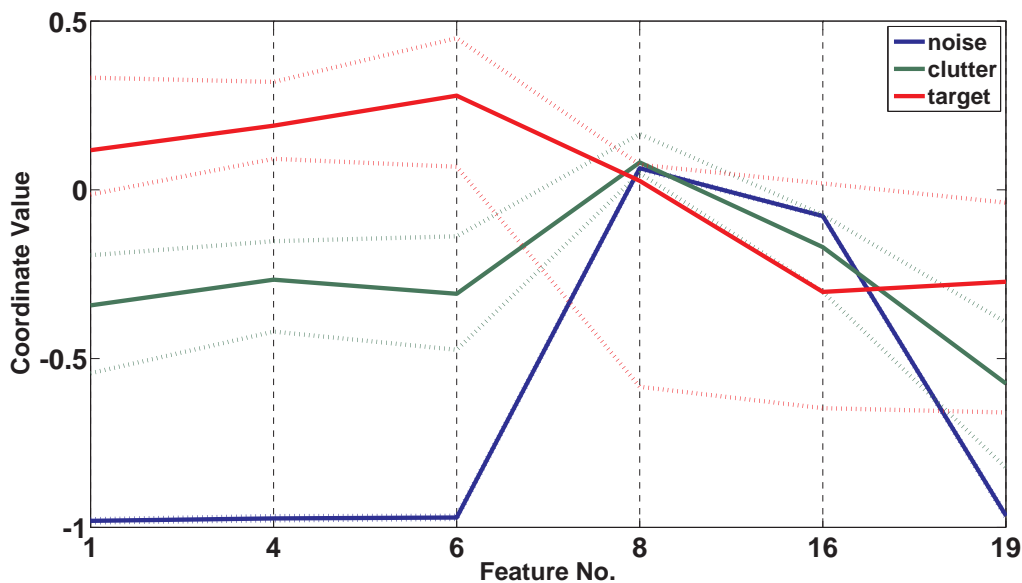
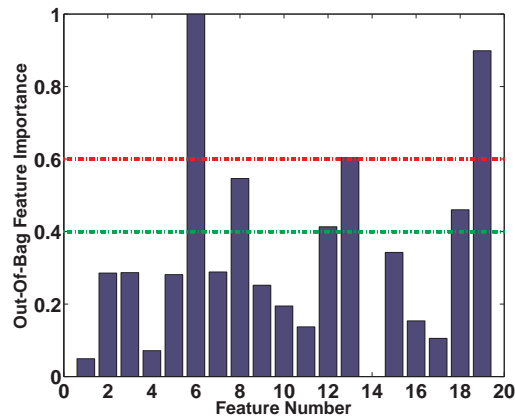


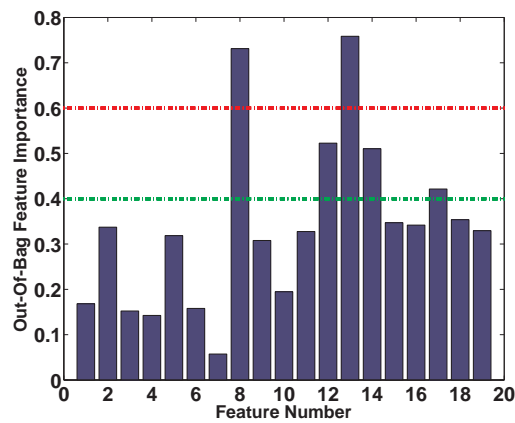
Figure 4.13. Parallel co-ordinates plot for the distance between the different classes using mixed training data.

Testing the different ensembles of trees which were trained using samples from the ‘one-target’, ‘two-target’ and ‘mixed’ data with both the ‘one-target’ data and the ‘two-target’ data. The test gives a classification error between 0 – 0.037 for matched cases, around 0.2 for the mismatched case, and between 0 and 0.0556 for the mixed training case. The classification error for all cases is listed in Table 4.5 in comparison to Mahalanobis distance classifier (using ICM and LSM segmentation methods).

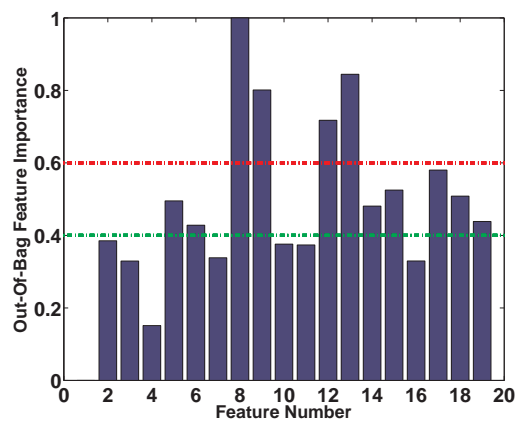
It is worth noting that the classification error is much higher in mismatched cases (where training is based on one-target data, and testing is used with two-target data and vice-versa), which implies a significant change in the features when other targets



(a)



(b)



(c)

Figure 4.14. Test for feature importance for classifying metal targets. The green and red line are guide-lines reflect a feature importance of 0.4 and 0.6 respectively. First, an ensemble trained on ‘one-target’ data is in (a), second, ensemble trained on ‘two-target’ data is shown in (b), and ensemble trained on mixed data in (c).

Table 4.4. Performance measures for the second ensemble using OOB samples, classifying metal target class

Training \ Measure	AUC	Optimal threshold	Optimal operating point
one target	1	0.589	(0, 1)
two targets	0.9995	0.6304	(0.027, 1)
mixed	0.9873	0.55	(0.0462, 0.9947)

Table 4.5. Classification error for the second ensemble. Classifying a metal target class. Segmentation by classification and Mahalanobis distance classifier (using ICM and LSM segmentation methods) are compared

		Seg. by Class.			ICM			LSM		
Testing \ Training	one target	two target	mixed	one target	two target	mixed	one target	two target	mixed	
	one target	0.037	0.2593	0.0556	0.6432	0.7584	0.7126	0.5335	0.8522	0.7126
two targets	0.2329	0	0	0.8334	0.6354	0.6635	0.7639	0.4698	0.5646	

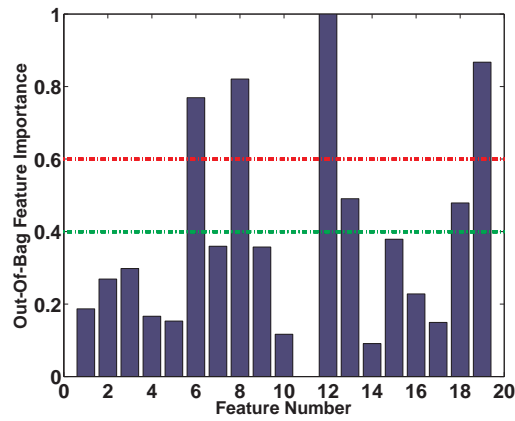
are present.

Figure 4.15 shows the feature importance when classifying target types. Table 4.6 lists the AUC, the optimal operating point and the optimal threshold for this case. Testing this case gives a classification error between 0 – 0.1429 for matched cases, between 0.3 – 0.94 for mismatched cases, and between 0 – 0.14 for the mixed training case. The classification error for all cases is listed in Table 4.7 in comparison to Mahalanobis distance classifier (using ICM and LSM segmentation methods).

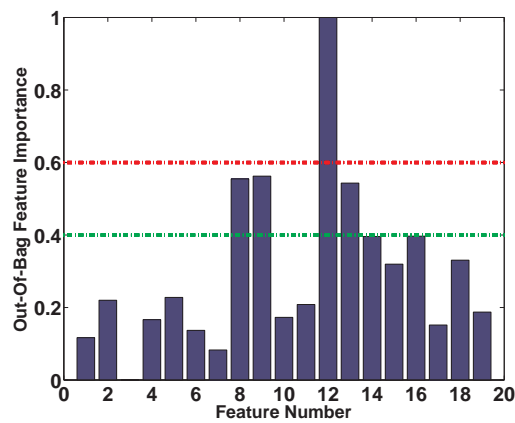
Table 4.8 lists the important features regarding each built ensemble, the most important feature has the symbol '■' and the features belonging to the six most important features are marked with '▲'.

4.5.4 General results

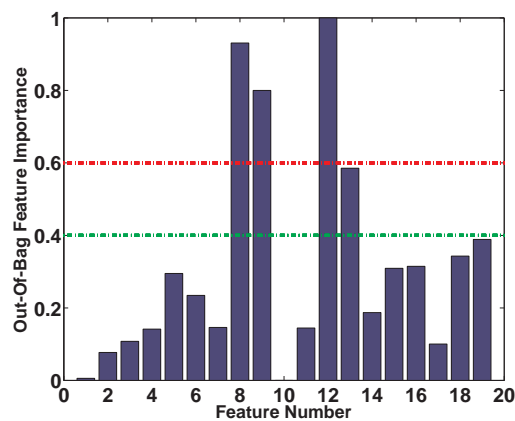
The features that were of general significance are (1, 4, 6, 8, 9, 12, 13, 19), which represent features from the main types mentioned in Section 3.2. It is also worth noting



(a)



(b)



(c)

Figure 4.15. Test for feature importance for classifying target types. The green and red line are guide-lines reflect a feature importance of 0.4 and 0.6 respectively. First, (a) is for ensemble trained on ‘one target’ data, (b) is for ensemble trained on ‘two-target’ data, and (c) is for ensemble trained on mixed data.

Table 4.6. Performance measures for the third ensemble using OOB samples, classifying target types classes.

Training \ Measure		AUC	Optimal threshold	Optimal operating point
dihedral	one target	0.9988	0.7917	(0, 0.9394)
	two targets	1	0.8333	(0, 1)
	mixed	0.9993	0.6579	(0.0073, 0.9915)
sphere	one target	0.9984	0.444	(0.013, 0.9697)
	two targets	0.9997	0.8286	(0, 0.96)
	mixed	0.9954	0.5714	(0.0152, 0.9828)
trihedral	one target	0.9970	0.7813	(0.01, 0.928)
	two targets	–	–	–
	mixed	0.9967	0.5135	(0, 0.9286)
saltwater jug	one target	0.9996	0.6857	(0, 0.9643)
	two targets	0.9988	0.675	(0, 0.973)
	mixed	0.9956	0.6071	(0, 0.9385)

that all of them are derived from HH and VV polarizations except for feature 19, the even bounce energy. Also, feature 8 (\log_{PHHV}) which is the logarithmic phase between HH and VV polarizations is an important feature in most of the trained ensembles. Most effective features for classifying between ‘target’, ‘clutter’, and ‘noise’ classes are (1, 4, 8, 19). For the metal target classification, the most important features are (6, 8, 12, 13, 15, 17), we see that features 15 (Im_{HHV^*}) and 17 (Im_{HV^*}) are excellent features in this case. They both involve the HV polarization, and they were not so important for ‘target’, ‘clutter’, and ‘noise’ classification. For the target type classification, we can see that feature 12 (Re_{HHV^*}) is the most important feature for all the trained ensembles on this case. Figure 4.16 and Figure 4.17 show the final segmentation and classification result for the first ensemble group that classified between ‘target’, ‘clutter’, and ‘noise’ classes using segmentation by classification compared to segmentation using ICM and LSM. The segmented scene

Table 4.7. Classification error for the third ensemble, Segmentation by classification, and Mahalanobis distance classifier (using ICM and LSM segmentation methods) are compared.

Testing \ Training		Seg. by Class.			ICM			LSM		
		one target	two target	mixed	one target	two target	mixed	one target	two target	mixed
dihedral	one target	0	0	0	0.352	0.432	0.214	0.354	0.514	0.5167
	two targets	0	0	0	0.626	0.3176	0.312	0.554	0.421	0.621
sphere	one target	0.1176	0.705	0.0588	0.475	0.8864	0.426	0.2841	0.878	0.3365
	two targets	0.3846	0	0	0.6276	0.3243	0.2116	0.5543	0.3236	0.4452
trihedral	one target	0	–	0	0.5447	–	0.3447	0.2332	–	0.2332
	two targets	–	–	–	–	–	–	–	–	–
saltwater jug	one target	0.1429	0.0714	0.1429	0.3632	0.4854	0.4362	0.2782	0.5321	0.4325
	two targets	0.9444	0	0	0.6387	0.3154	0.2543	0.8854	0.366	0.4265
overall	one target	0.0741	0.3704	0.0556	0.4337	0.6017	0.355	0.2873	0.6417	0.3795
	two targets	0.3014	0	0	0.6307	0.3191	0.2593	0.6645	0.3702	0.4977

contained two metal dihedrals at the same distance from the platform and at different heights. It can be seen from the figure that using simple segmentation methods depending only on pixel intensity is able to get the real positions of the targets, but it suffers from ghost targets as in the case of ICM or from large false positive detections as in the case of LSM. Taking into account that the three algorithms have the same initial pixel intensity threshold to begin with, segmentation by classification leads to better segmentation results than the other algorithms.

4.6 Conclusions

The problem of target segmentation and classification in the image-domain with application to Through-the-Wall Radar Imaging has been addressed. Perceptually meaningful entities, obtained from a low-level grouping process, the so-called super-pixels, are considered in this thesis, overcoming the limitation of pixel-grid based

algorithms. Further, simple geometrical and statistical descriptors that have been used as features in previous work are extended to polarimetric descriptors, which make use of the whole polarimetric information in radar images. A framework of polarimetric feature extraction, over-segmentation (super-pixels), clustering, and classification has been presented. An expandable sequential classifier based on random forests has been proposed to discriminate targets from clutter returns and to provide further information about the discriminated targets. The experimental results demonstrate the usefulness of the proposed methods as desired target returns are discriminable from clutter returns and a further classification about target type and its nature is provided. The proposed technique proved better performance than techniques that depend only on pixel intensity.

Table 4.8. Feature importance for each built ensemble

		features																		
Classify	Built on	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Target or not	1 target	▲		▲	■		▲				▲					▲				
	2 targets	▲			■				▲	▲					▲					▲
	mixed	▲			■		▲		▲								▲			▲
Metal or not	1 target						■		▲				▲	▲		▲				▲
	2 targets								▲				▲	■	▲			▲	▲	
	mixed								■	▲			▲	▲		▲		▲		
Target types	1 target						▲		▲				■	▲					▲	▲
	2 targets								▲	▲			■	▲	▲		▲			
	mixed								▲	▲			■	▲					▲	▲

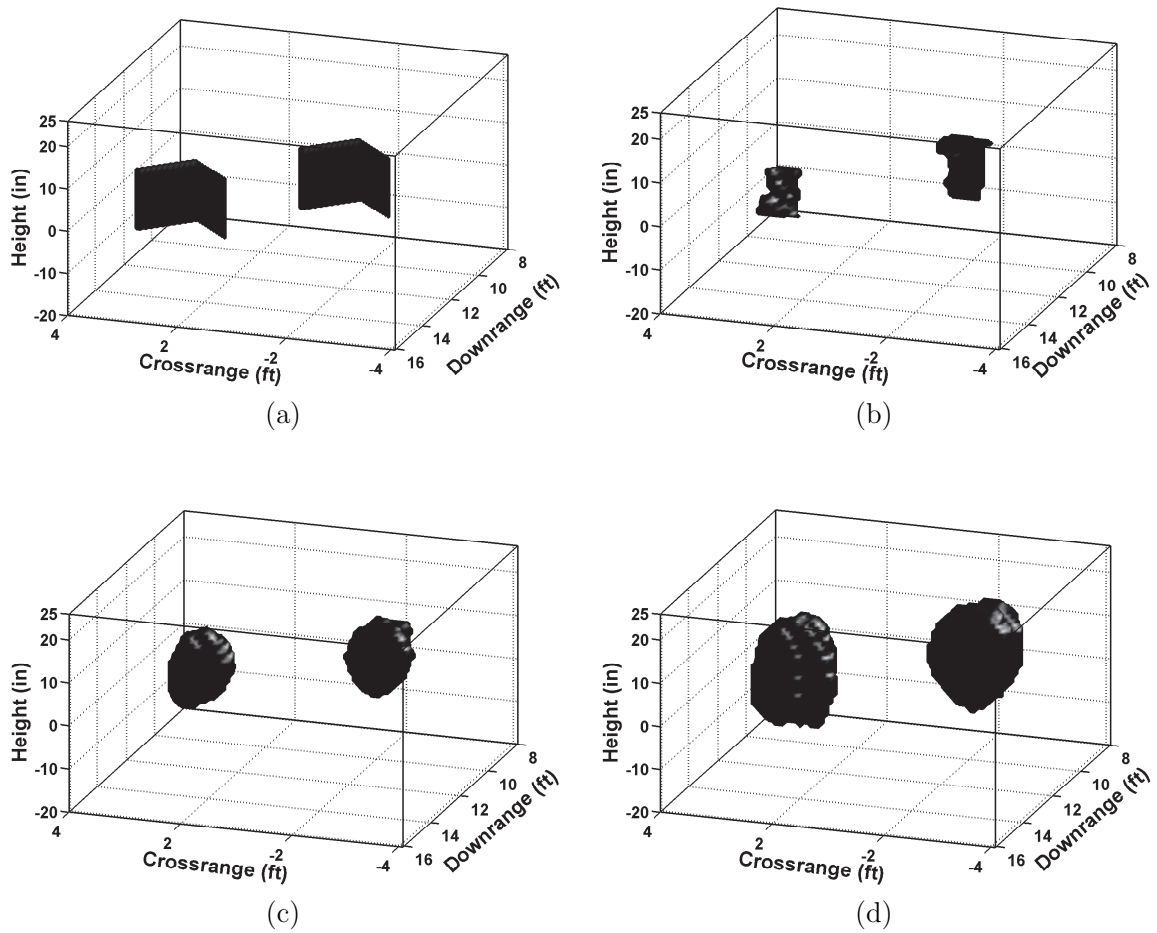


Figure 4.16. Front view of final 3D segmentation of a two target scene containing 2 dihedrals, (a)Original scene (b) Segmentation by Classification (c) ICM (d) LSM.

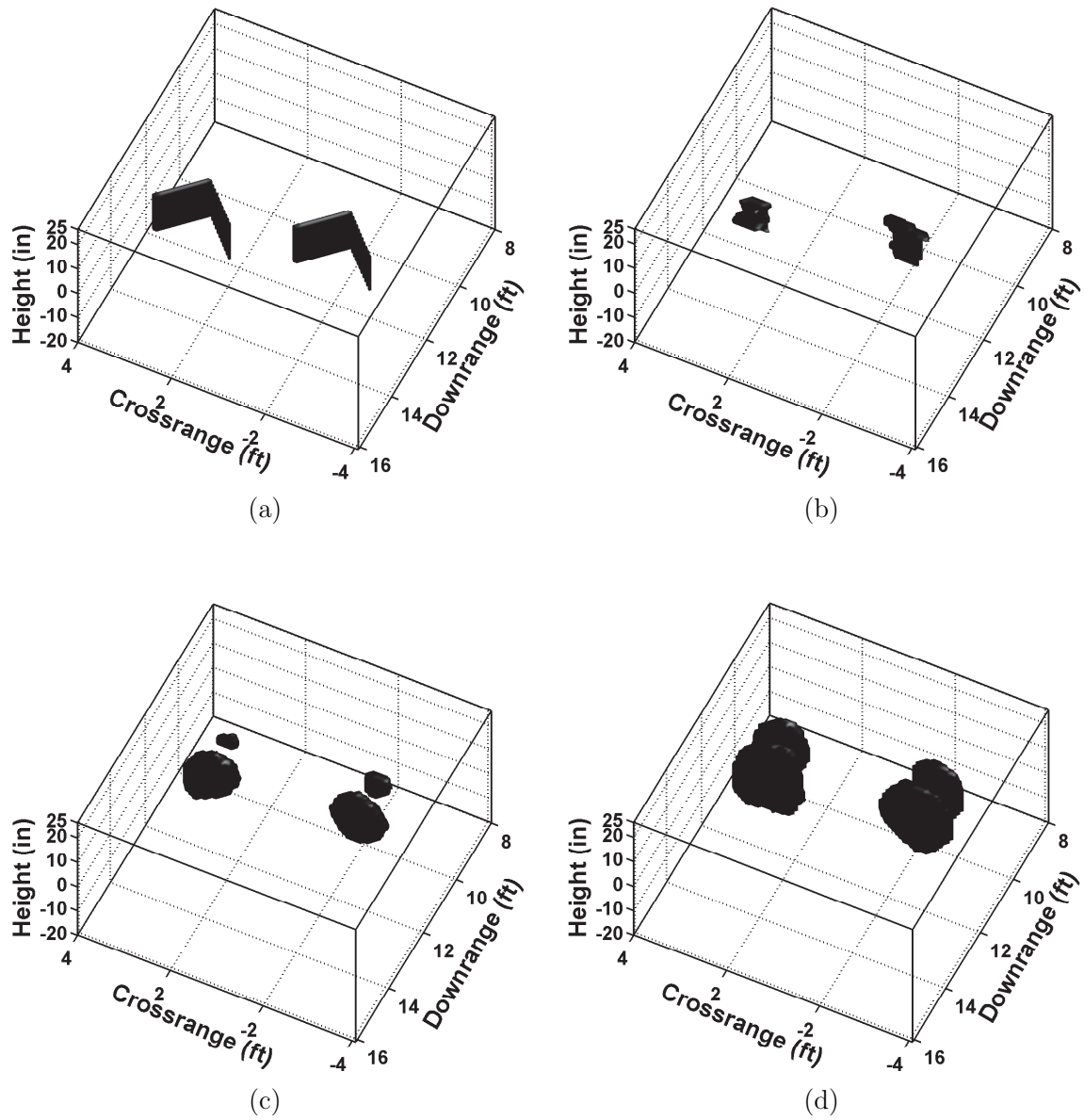


Figure 4.17. Top view of final 3D segmentation of a two target scene containing 2 dihedrals, (a)Original scene (b) Segmentation by Classification (c) ICM (d) LSM.

Chapter 5

Multi-level 3D classification

In this chapter, 3D classification of stationary targets in the image domain is considered. The TWRI image is divided into a finite set of segmented regions or objects. Each of these objects is labeled to a certain class according to the values of the different extracted features.

Research on TWRI target detection has been developed in two domains, the data-domain and the image-domain. Data-domain target detection involves waveform design to improve target detection. This exploits matched illumination with partial or full prior knowledge of the target Radar Cross Section (RCS) over angle and frequency [32, 134–136]. These techniques work well under specific assumptions about the targets and propagation environments. With multiple targets and unknown target orientation, these techniques become less effective. Image-domain based target detection [7, 54, 73, 75, 76, 84, 87] handles multiple targets with no prior assumptions on the target RCS. These techniques face the challenge of operating with limited bandwidth and insufficient physical or synthesized array aperture, thus disallowing high resolution based target analysis and classification. Detection in the image-domain has been proposed using centralized [7] as well as decentralized [74] approaches. The common aim within both approaches is to deduce a single binary, three-dimensional, reference image from a set of 3D TWRI images, to detect the presence or absence of targets. Most of these approaches make use of single polarization data while multi-polarization data can capture different aspects of a target and is, thus, promising to improve the overall detection result [87, 137, 138].

Existing work in polarimetric imaging for TWRI applications includes [10, 14, 80, 87, 139]. In [10], cross- to co-polarization return ratio was used for rifle detection. In [139], the polarization difference imaging is employed to improve enhancement in terms of image quality. In [14], the effect of ghost targets, that appear in TWRI images due to multi-path propagation is reduced using a multiplicative combination technique of co- and cross-polarized time-difference images. As such, the image quality is strongly improved which facilitates target detection. In [85], the performance of the adaptive target detection techniques proposed in [7, 73, 76] is improved utilizing the polarization diversity and accounting for changing image intensity distributions as a function of transmitter-receiver co- and cross-polarizations

In the previous Chapter, target segmentation and classification scheme consisting of feature extraction, segmentation and classification in TWRI applications has been

introduced. The scheme involves super-pixels instead of the regular pixel-by-pixel grid representation. The motivation was to use a more natural representation that involves homogeneous characteristics of the segmented regions. In this Chapter, two new schemes are proposed. The first involves classification using estimated probabilities of co-occurrence matrices (EPCM) and fusion-based log likelihood ratio test. The second involves texture features based on co-occurrence matrices and polarimetric features (PolTex) combined with state-of-the-art classifiers. The polarimetric signatures in the second framework are the features used in the previous Chapter in addition to new intensity spatial features. Discrimination between target / non-target classes followed by target type classification is demonstrated. The organization of this Chapter is as follows. Section 5.2 reviews the features that will be used throughout this chapter that were introduced in Chapter 3. Section 5.3 presents the clustering phase, where similar voxels are grouped into clusters, and the unwanted clusters are excluded. Section 5.4 discusses the different schemes for classifying voxels of the remaining cluster into ‘target’, ‘clutter and noise’ voxels. Section 5.5 presents experimental results, and conclusions are drawn in Section 5.6.

5.1 Motivation

Target detection in TWRI is a computationally expensive process, due to the large number of pixels or voxels under test. TWRI applications usually requires fast results as it mostly deals with life and death situations as in search and rescue operations, law enforcement, and military applications. To reduce the computational time, a multi-step approach is needed. The first step uses k-means clustering [140] exploiting spatial polarization features to produce a reduced set of voxels. This is achieved by excluding the clusters that mostly contain noise and clutter voxels which will result in a preliminary view about target positions. The second step involves calculating polarization signatures and co-occurrence matrix features for this reduced set of voxels. These features are used in different arrangements to classify the voxels to target / non-target (clutter and noise) classes. The contribution in this chapter is a framework for 3D target detection using polarization signatures and co-occurrence matrix features. A diagram summarizing the general steps of the proposed algorithms is shown in Figure 5.1.

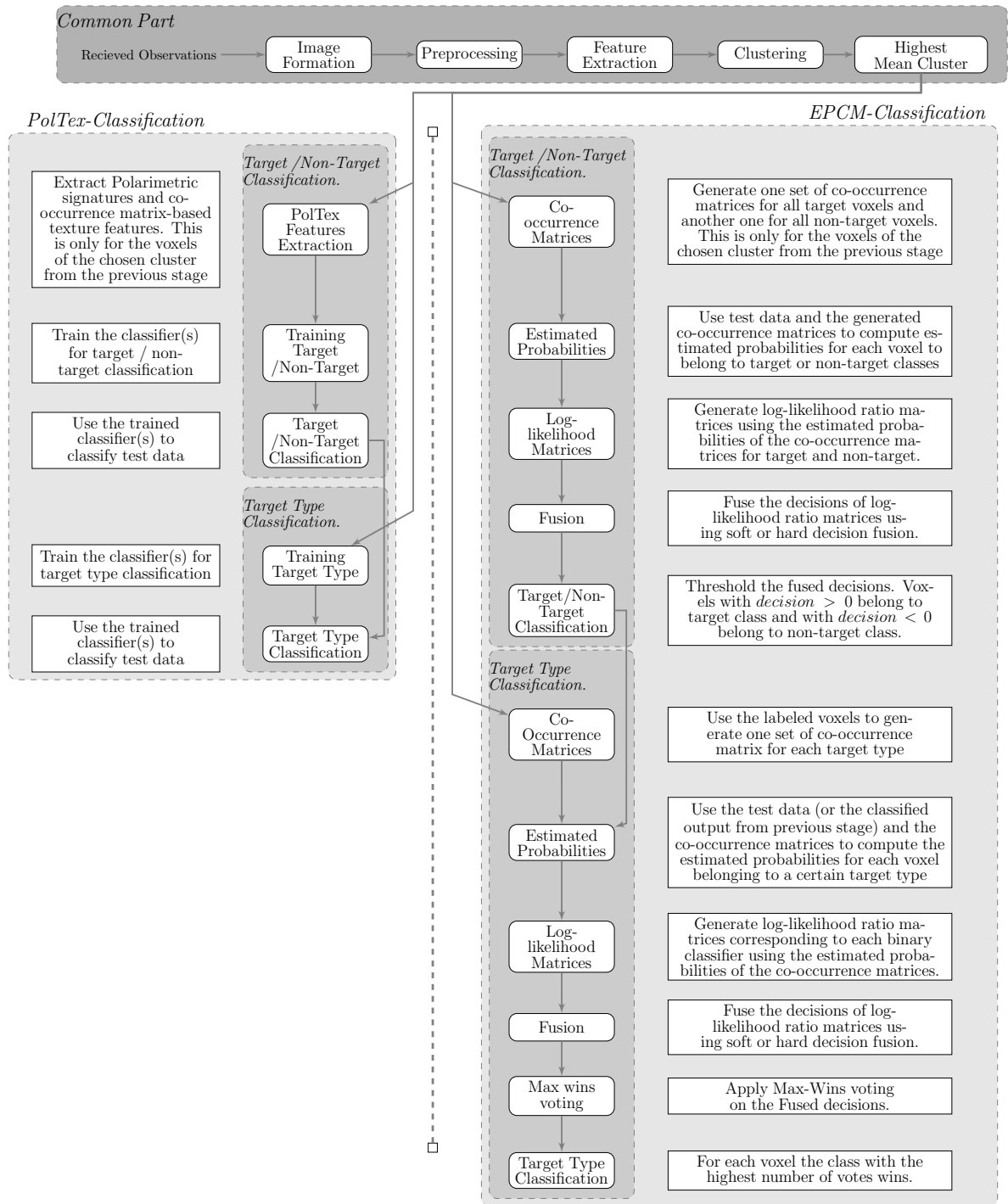


Figure 5.1. The general steps of the proposed frameworks.

5.2 Feature Extraction

As explained in Chapter 3, different sets of features are used throughout the proposed algorithms, in Chapter 4 polarimetric signatures were used as features in the

segmentation by classification algorithm. In this chapter, features are used in a multi-stage approach to reduce computational complexity and time. In Section 5.3, the spatial polarimetric signatures from Section 3.2 are exploited to remove most of the noise and clutter in the image. In Section 5.4.1, the co-occurrence matrices from Section 3.3 are exploited to obtain estimated probabilities matrices for the different classes in test. In Section 5.4.2, polarimetric signatures from Section 3.2 and texture features from Section 3.3 are both used as features for classification. The two algorithms mentioned only extract the features from the regions remaining after the clustering phase. To eliminate the dependence of the features on target resolution and location, the algorithm presented in [85] can be implemented. For simplicity, we use the features without compensation.

5.3 Clustering

We aim at clustering the voxels to a maximum of four non-overlapping regions. The four clusters represent 2 noise classes, one clutter class, and one target class. Two noise classes are assumed due to the high number of noise sources [8, 9, 13, 141] as mentioned in Chapter 1. We use k -means clustering [140] as it is a simple and fast algorithm. It has comparable performance to more complex algorithms, and works well for compact and hyper-spherical clusters. Further, it can be used to cluster large data sets [142, 143]. The feature vector used as an input to the clustering algorithm exploits thresholded intensities and the spatial intensity features. These features are chosen because they are computationally inexpensive and provide spatial information. Referring to Table 3.1, the chosen features are number 1-3 and 20-28. Hence, the feature vector of each voxel has 12 features. Figure 5.2 shows a scatter plot of the clustered voxels in a space of the normalized median of the HH and HV intensities. Figure 5.3 shows the density function of the four clusters with the mean of each cluster denoted by a dashed vertical line. The target cluster has the highest mean, and this is how it is identified in our algorithm. Figure 5.4 shows the chosen cluster in a 3D scene in black. The green cube is a bounding box for the original target place. The cluster contains most of the target voxels in addition to clutter and ghost voxels. The clusters that are expected to contain only noise and clutter are excluded. Only a negligible part (less than 1%) of target voxels is removed, if any. We will continue with the highest-mean clusters that are assumed to contain most of the target voxels in the following experiments. This yields a preliminary estimation of the target position in a very short time and provides the possibility to do more extensive computations on the remaining voxels.

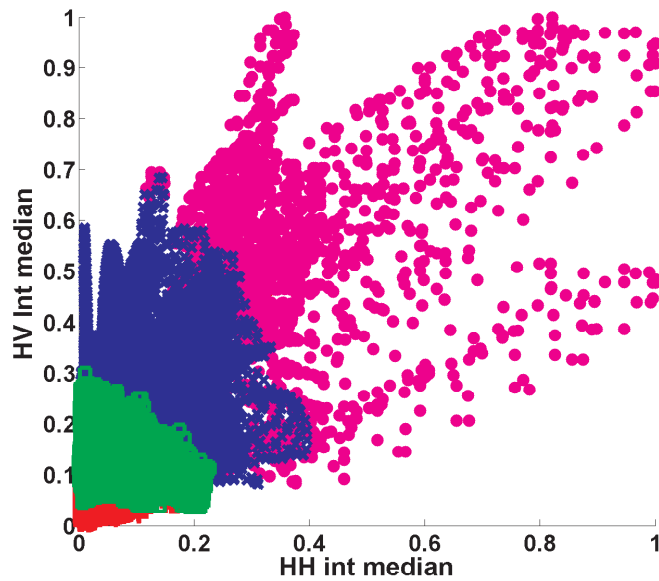


Figure 5.2. Clustered voxels for a scene with one target, cluster with the higher intensity values contain target voxels (magenta). Other clusters contain clutter (dark blue) and noise (green and red)

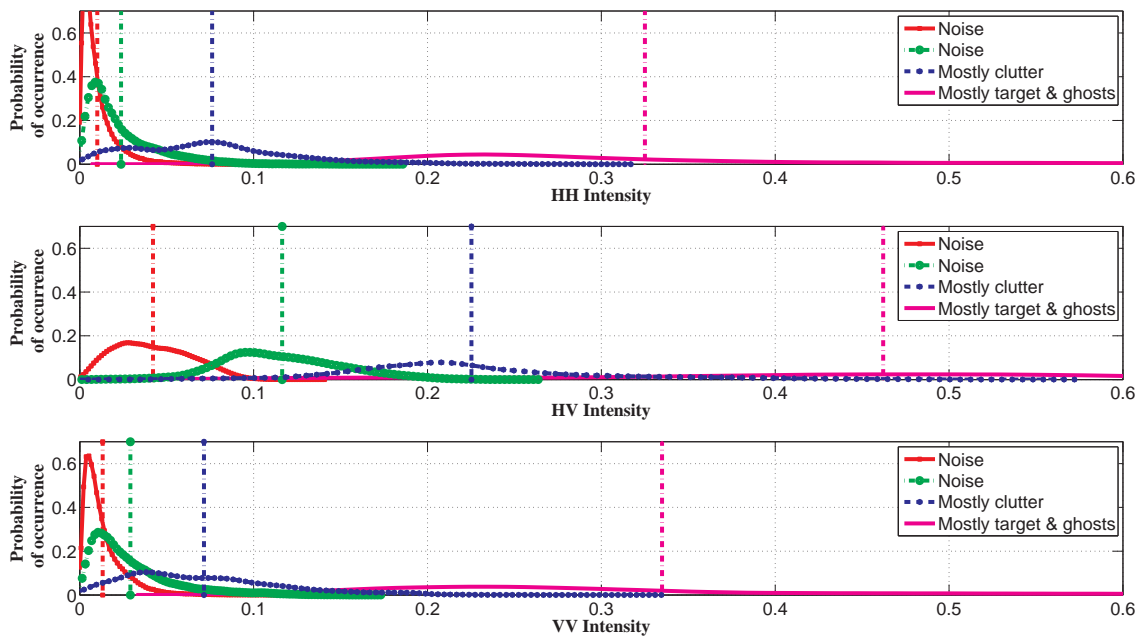


Figure 5.3. Empirical density function for the four clusters of a one-target scene

5.4 Classification

In this section, we describe how to further classify the group of voxels resulting from the clustering stage to get a better representation of the target. In [62, 86],

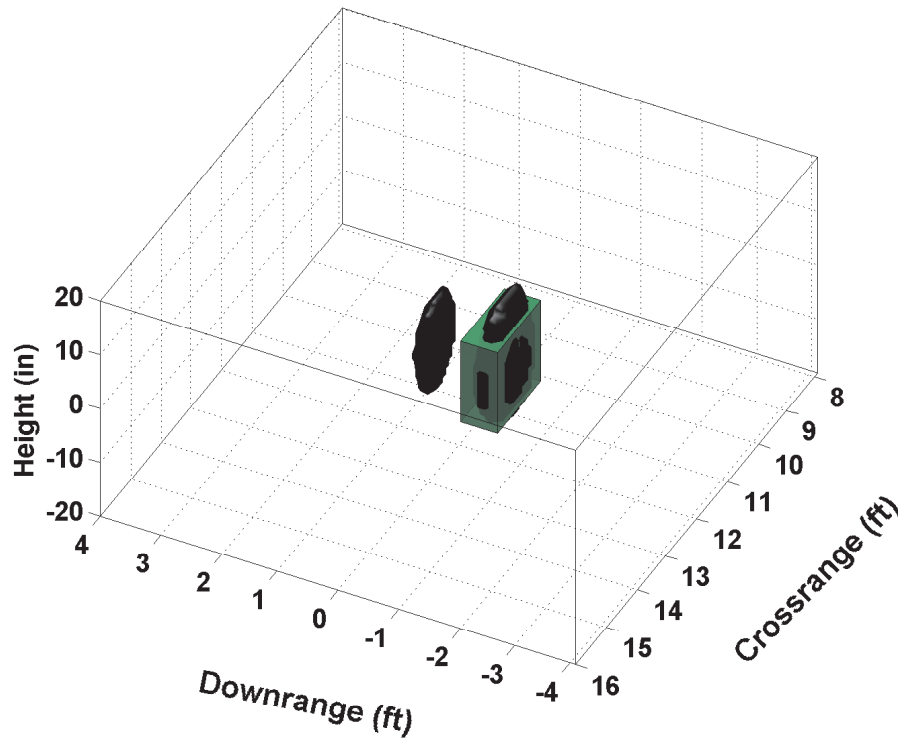


Figure 5.4. Chosen cluster in a 3D scene in black. The green cube is a bounding box for the original target place. This shows the arrangement of experiment 1 with one target (dihedral)

the minimum Mahalanobis distance classifier is used to classify targets exploiting 3D statistical models in stationary target classification for TWRI. These models depend only on HH single channel polarization and do not provide a possibility to handle the resolution change. Further, in [84], Iterative Conditional Modes (ICM) segmentation was considered to divide the TWRI image into target and noise, then the nearest-neighbor classifier and support vector machines were used for further classification, employing super-quadric features. This technique assumes that the pdf of the target is known and has no measure of segmentation phase performance, although the classification phase depends on the volume feature of the segmentation results. In [85], the probability density function (pdf) of the image pixels attributed to the target is estimated and super-ellipsoids are fitted to imaged targets. The pdf and the super-ellipsoid parameters were combined into a feature vector and passed to the classifier. The technique compensates for the variation of system point spread function (PSF) over the image, which is of practical importance. This technique extracts weibull model parameters from target candidates which may include clutter and noise. Also, the super-ellipsoid is highly non-linear and shape fitting can be complicated, com-

putationally expensive, and problematic [144] especially with partial or insufficient data. In [87], a scheme for target detection using segmentation by classification is proposed. The scheme exploits polarimetric signatures to create super-pixels and then classify them using random forest decision trees. This technique has comparable performance to the aforementioned algorithms, but it has the disadvantage of averaging the features throughout the whole super-pixel. To avoid feature averaging among different voxels, we propose two new algorithms that exploit the neighborhood voxels in feature extraction and either use the features directly or through a feature fusion algorithm. Each of the two classification schemes can be used in different configurations. The first classification algorithm uses only the co-occurrence matrix estimated probabilities and performs fusion-based log-likelihood-ratio test for classification. The second method uses the polarimetric signatures and the texture features from the co-occurrence matrix, combined with state-of-the-art classifiers.

5.4.1 Classification using the estimated probabilities of co-occurrence matrices (EPCM)

5.4.1.1 Target / non-target classification

As mentioned in Section 5.3, we begin with the highest-mean cluster. The target is to label voxels as target / non-target voxels. For the training part, two different sets of co-occurrence matrices are generated. One set is built using the target voxels and the other using the non-target (clutter and noise) voxels. An important parameter at calculating the co-occurrence matrix is the number of gray levels used to represent the image. Having more gray levels will better represent the variation of intensities in the image, especially in our case where small variations are present. On the other hand, having more gray levels will result in sparse co-occurrence matrices especially in higher dimensions. Tests with $G \in \{2, 4, \dots, 16\}$ are conducted to choose the optimum number of gray levels. Each test has 54 different clique arrangements as illustrated earlier in Figure 3.3. This results in 432 different co-occurrence matrices for each set. Note that this high number is just for finding the optimum number of gray levels. Once it is determined, we will have only 54 configurations. Let $\Omega = \{\omega_0, \omega_1, \dots, \omega_M\}$ be the set of classes, where M is the total number of classes. Given a class ω_m , the co-occurrence matrix generated using the clique φ_r is denoted as $C_{\varphi_r}(\mathbf{i}|\omega_m)$, where $\mathbf{i} = [i_1, i_2, \dots, i_k]$. In the target / non-target case $M = 2$, so we will have 2 sets of co-occurrence matrices for target and non-target. These two sets of co-occurrence matrices result from the training phase as shown in Figure 5.5.

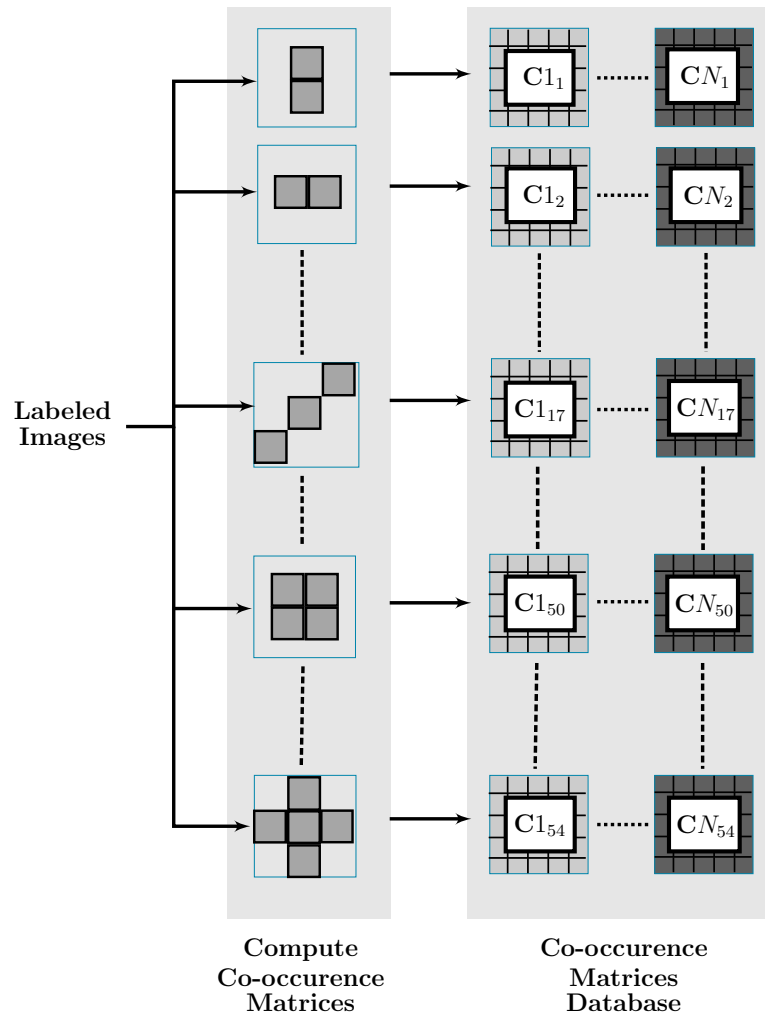


Figure 5.5. Training steps to get co-occurrence matrices

The testing phase involves calculating the estimated probabilities using the previously generated co-occurrence matrices. Figure 5.6 shows a sample co-occurrence matrix estimated probabilities generated from the training phase. The process of generating the estimated probabilities of each voxel in a sample image using the co-occurrence matrix from Figure 5.6 is shown in Figure 5.7. The same MRF cliques are used to estimate the probability matrices. Note that the estimated probabilities in the figure are just an example and do not represent any real data. A set of all configurations of a clique φ_r in a lattice \mathcal{L} can be defined as $\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_N\}$, where γ_n denotes the indices representing a clique at configuration n , and N is total number of configurations a clique can have in a test image. For each voxel q , the estimated probability of this voxel belonging to class ω_m using clique φ_r is computed by accumulating the relative frequency values for each voxel as

$$\hat{p}_{\varphi_r}(q|\omega_m) = \begin{cases} \sum_{n=1}^N \mathbf{1}\{q \in \gamma_n\} \cdot C_{\varphi_r}(g(\gamma_n)|\omega_m), \\ 0, \text{ otherwise} \end{cases} \quad (5.1)$$

where $g(\gamma_n)$ represent the gray levels of the voxels with indices denoted by the set γ_n .

Given the probability estimates, log-likelihood ratios between the two classes ω_0 and ω_1 can be calculated per clique as

$$L_{\varphi_r}(q) = \log \frac{\hat{p}_{\varphi_r}(q|\omega_1)}{\hat{p}_{\varphi_r}(q|\omega_0)} \quad (5.2)$$

The reliability information measured by $|L_{\varphi_r}(q)|$, and the decision rule, $\text{sign}(L_{\varphi_r}(q))$, specify the likelihood ratio test completely. For a more reliable decision, a decision fusion scheme can be employed, especially if there are a large number of likelihood ratios from different sources. In MRFs, the random variables \mathcal{X} corresponding to different cliques are independent. Therefore, the optimum fusion rule is equivalent to the soft decision fusion with R independent sources [145]

$$L(q) = \sum_{\forall r} L_{\varphi_r}(q) \underset{\omega_0}{\overset{\omega_1}{\gtrless}} \log \tau, \quad (5.3)$$

where τ is the constant threshold, and in our case, it is set to 1. The reliability information per clique can be misleading, incorrect, due to an incorrect estimation of $\hat{p}_{\varphi_r}(q)$, or affected by an uncertain event, such as impulsive noise. In such cases, hard decision fusion can be more suitable because the likelihood ratios are thresholded and the effect of very large and unreliable L_{φ_r} in equation (5.2) is circumvented. The optimum hard decision fusion is defined as the weighted sum of decisions of each information source [145]

$$\sum_{r=1}^R u_r \log \frac{1 - P_{Mr}}{P_{Fr}} + (1 - u_r) \log \frac{P_{Mr}}{1 - P_{Fr}} \underset{u_r=0}{\overset{u_r=1}{\gtrless}} \log \tau, \quad (5.4)$$

where P_{Fr} and P_{Mr} are the probabilities of false alarm and miss detection for the corresponding clique φ_r , and τ is a suitable threshold as mentioned earlier. In our problem, the information source is parameterized by φ_r , corresponding to the decision u_r . Let $\log \frac{1-P_{Mr}}{P_{Fr}} = \alpha_r$ and $\log \frac{P_{Mr}}{1-P_{Fr}} = \beta_r$. The setting of $\alpha_r = 1$ and $\beta_r = -1$ correspond to the hard decision fusion stated in our experiments. However, α_r and β_r are not necessarily equal to 1 and -1, respectively. They can be also estimated

	0	1	2	3
0	0.15	0.03	0	0.03
1	0.07	0.25	0.1	0.15
2	0	0.04	0.12	0
3	0	0	0.04	0.02

Figure 5.6. A sample co-occurrence matrix for a 4 gray level image, generated in the training phase

0	1	0	3	2	1	0.03	0.03	0.0	0.0	0.0	0.0
1	1	2	3	1	1	0.0	0.0	0.0	0.0	0.0	0.0
3	2	0	1	2	2	0.0	0.0	0.0	0.0	0.0	0.0
2	3	3	2	1	0	0.0	0.0	0.0	0.0	0.0	0.0
(a)						(b)					
0	1	0	3	2	1	0.03	0.1	0.07	0.0	0.0	0.0
1	1	2	3	1	1	0.0	0.0	0.0	0.0	0.0	0.0
3	2	0	1	2	2	0.0	0.0	0.0	0.0	0.0	0.0
2	3	3	2	1	0	0.0	0.0	0.0	0.0	0.0	0.0
(c)						(d)					
0	1	0	3	2	1	0.03	0.1	0.1	0.07	0.08	0.04
1	1	2	3	1	1	0.25	0.35	0.1	0.0	0.0	0.25
3	2	0	1	2	2	0.04	0.04	0.13	0.2	0.22	0.1
2	3	3	2	1	0	0.0	0.02	0.06	0.08	0.11	0.07
(e)						(f)					

Figure 5.7. Sample for generating the estimated probabilities matrix in the testing phase (based on the sample co-occurrence matrix from Figure 5.6) for a 4 x 6 image. The left column show the movement of the clique through the test image, while the right column show the corresponding change in the estimated probabilities matrix. The estimated relative frequency corresponding to the current 2 gray levels marked by the clique, is added to the 2 corresponding cells in the estimated probabilities matrix. A bivariate clique is used with distance vector $\Lambda = (1, 0)$.

directly from the likelihood ratios at the training phase to benefit from the soft decisions. Any likelihood ratio based on a clique, which is unreliable or does not contribute to the detection performance will be assigned a zero weight. To this end, we use a sequential feature selection (SFS) algorithm [146] to select suitable likelihood ratios for our problem. The selection procedure can be improved upon

determining the weights with a better precision. Instead of 0 and 1, each likelihood ratio can be allowed to be multiplied by a real number κ , e.g $\kappa \in \{0.1, 0.2, \dots, 1\}$. Higher precision clique selection has very minor improvement. Therefore, we constrain ourselves only to the binary weights with SFS. A block diagram explaining the testing process is shown in Figure 5.8. In the case of target / non-target classification we have only two classes, so only the first row of calculating the log-likelihood phase will be implemented.

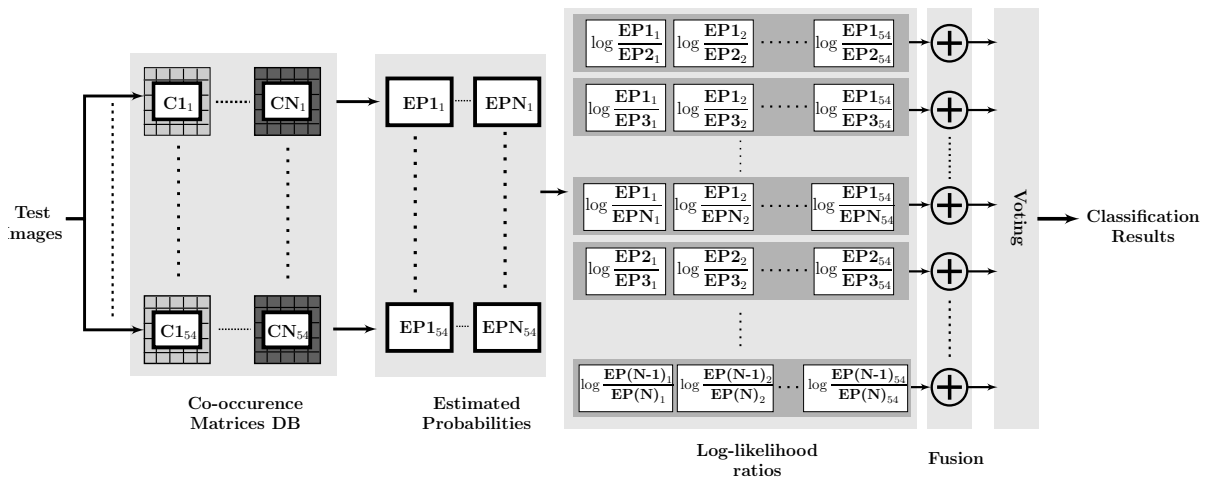


Figure 5.8. Testing phase to get classification results

5.4.1.2 Multi-class Classification

In order to classify the voxels to multiple target types, we extend the previous algorithm to multi-class classification. In this case, we use a voting algorithm to assign each voxel to its class. The max-wins algorithm [147–149] is chosen for its low computational complexity and good performance. Assume we have M classes with labels w_i where $i = 1, 2, \dots, M$. The max-wins method involves creating a binary classifier for every distinct pair of classes. This sums to $C = M(M - 1)/2$ binary classifiers. We used the log-likelihood from equation (5.2) as our binary classifier between each pair of classes w_i and w_j . A set of log-likelihood matrices corresponding to the 54 clique arrangements is generated for each class exactly as in the target / non-target case. The test will run exactly in the same configuration for each binary classifier. Soft and hard decision fusion is performed for the output decisions of all the clique arrangements. For each voxel on the test image, a vote is given that this voxel belongs to class w_i if the value on the decision matrix is positive and to class w_j if the value is negative. For each voxel, there will be C final

decisions or votes. The voxel is assigned to the class with the highest number of votes. In case two classes have the same number of votes, the vote of the classifier comparing these two classes is used to predict the class. A block diagram explaining the training and testing processes is shown in Figure 5.5 and Figure 5.8.

5.4.2 Classification Using Polarimetric Signatures and Texture Features (PolTex)

In this part, the polarimetric signatures and texture features extracted from the co-occurrence matrices are used. Each voxel has a vector of 64 features (28 polarimetric features and 36 texture features). In this classification problem, we are given a set of N voxels for training, $\chi = \{(\mathbf{F}_1, w_1), \dots, (\mathbf{F}_N, w_M)\}$, where \mathbf{F}_n is a vector of descriptors and w_m is the corresponding m -th class label. Our goal is to find a model for predicting the class w from new feature vector \mathbf{F} . For simplicity, we will stick to the optimum number of gray levels chosen at the previous stage to compute the texture features. Four classifiers, which include the Linear Bayes Normal Classifier [126, 150], Quadratic Bayes Normal Classifier [126], K-Nearest Neighbor Classifier [126, 151], and Fisher Linear Discriminant Classifier [152, 153] are used for target / non-target classification and target type classification. An SFS algorithm is applied to choose the optimum feature set for each classifier. As the optimum set of features may vary in size, we test the performance of a small number of features through all classifiers to compare the performance. In our experiments, classification performance is compared using the best ten features for each classifier.

5.5 Experimental results

The experimental setup is depicted in Figure 2.4. We consider the scenario presented in Chapter 2 for the evaluation of the proposed techniques. As mentioned, we have four *one-target* 3D images, two *two-target*, and one *four-target* 3D image. Each 3D image is constructed from a set of 57×57 2D B-Scans in three different polarizations, HH, HV and VV resulting in 21 different sets.

At first, polarimetric intensity and spatial signatures for the images are extracted. Each voxel will have a vector of 12 features. It is well known that the k-means clustering method is sensitive to outliers [154, 155]. The set of feature vectors should be checked for outliers (very large values in amplitude) and scaled so as to avoid an

undesired behavior of the clustering step, such as deviated clusters center or clusters that contain only an outlier. The images are then clustered to four non-overlapping regions. The cluster with the highest mean is chosen as an appropriate candidate for further classification. In the training and testing phases, the used voxels from the database should be labeled to target and non-target voxels. Also for further classification, the target voxels are labeled according to target types.

We have two different sets of classification experiments. The first set involves classification using the estimated probabilities of the co-occurrence matrices (EPCM) and the second set is exploiting polaremitric signatures and texture features based on co-occurrence matrices (PolTex).

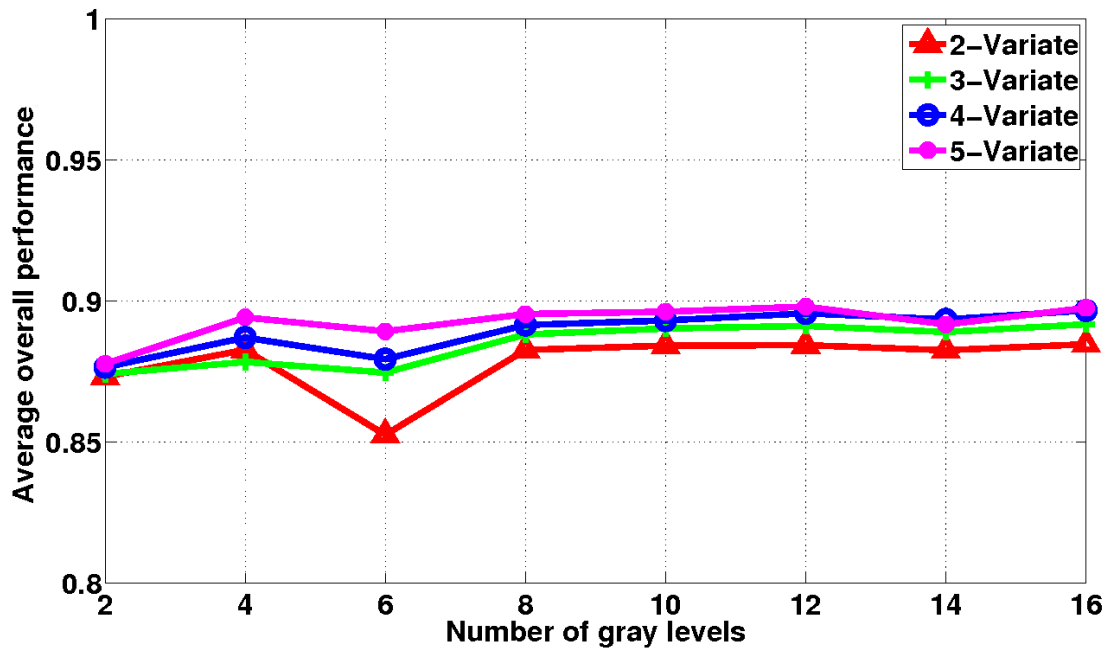
5.5.1 Classification experiments using the co-occurrence matrix estimated probabilities

5.5.1.1 Experiments description

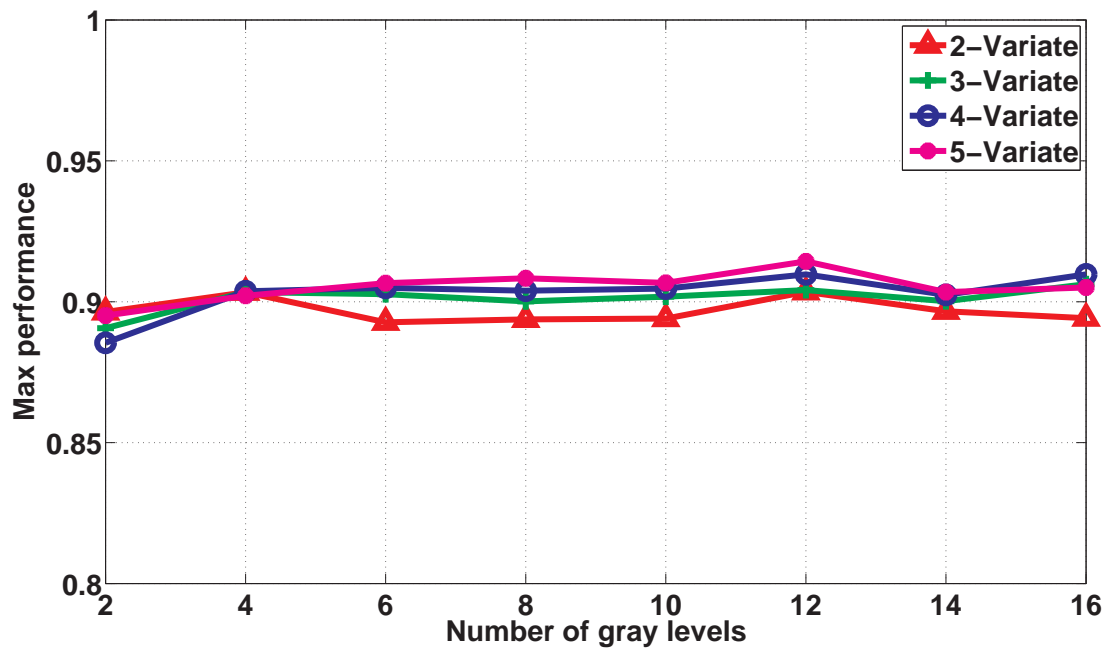
This set of experiments are arranged following the EPCM classification part of the scheme shown in Figure 5.1. To choose the optimum number of gray levels, eight different numbers of gray levels are tested. The 54 clique arrangements in the set Φ are used with each number of gray levels tested. The experiments are done in a 5-fold cross-validation arrangement, the original dataset is randomly partitioned into 5 subsamples. All of the subsamples are equal in size. A single subsample is used for testing, and the 4 subsamples remaining are put together and used as training data. The process is repeated 5 times, with each of the 5 subsamples used exactly once as the testing data. This yields 5 results from the folds, which are averaged to produce a single estimation. The advantage of this method over repeated random sub-sampling is that all observations are used for both training and testing, and each observation is used for testing exactly once.

5.5.1.2 Experiments results: optimum gray level

The results of a 5-fold cross-validation implementation of these experiments show that using 12 gray levels gives generally the highest performance. Figure 5.9 shows both the average and the maximum performances for different clique dimensions through different gray levels. In Figure 5.9, we can see that performance differs slightly between the levels from 8 to 16. The following results are obtained using images quantized at 12 gray levels.



(a) Overall average classification performance



(b) max classification performance

Figure 5.9. Average and maximum performance for different clique dimensions against the number of gray levels.

5.5.1.3 Experiments results: classification

The results for both target / non-target classification and target type classification are shown in Table 5.1. Soft and hard decision fusion, forward selection and the

best 10 arrangements are used to handle the different log-likelihood outputs corresponding to the 54 cliques employed. The performance is compared to the Linear Bayes Normal Classifier (LBN), the Quadratic Bayes Normal Classifier (QBN), the K-Nearest Neighbor Classifier (KNN), and Fisher's Least Square Linear Classifier (FLS) using thresholded HH intensity (first feature in polarimetric signatures). All classifiers use the output voxels from the clustering stage, and the results are based on a 5-fold cross-validation. For target types, we have four classes: dihedral, sphere, trihedral, and a salt water jug. Classification using co-occurrence matrices for target / non-target classification has lower error rates compared to the other classifiers. Although using our method for classifying target types yields lower error rates compared to the other tested classifiers, the overall performance is low. This performance deterioration is mostly caused by the small amount of samples available to fill the multi-variate co-occurrence space, where the samples available for each target type is about 0.25 of the samples available to the 'target' class. This can lead to sparse co-occurrence matrices, where many elements are zero, which in turn does not represent the differences between the classes. Table 5.2 shows the class error rates for both target / non-target and target type classification.

Table 5.1. Overall classification error rates for the target / non-target and target types classification compared to other types of classification algorithms using only HH polarimetric signature as a feature.

	Co-oc matrices.				Compared classifiers			
	Soft Fusion	Hard Fusion	Opt. Set	Best 10	LNB	QNB	Fisher	KNN
Target / non-target	0.0728	0.0784	0.0638	0.0648	0.1697	0.181	0.175	0.163
Target types	0.389	0.382	0.380	0.383	0.4516	0.4606	0.4511	0.4315

As mentioned earlier, a forward selection algorithm is used to choose the optimum set of features for classification. Optimum performance was achieved by 27 clique arrangements in the target / non-target classification and 29 clique arrangement in target type classification. Clique numbers can be viewed in Figure 3.3. Table 5.4 lists the contribution of each clique dimension (considering the number of elements in a clique) to the optimum set.

From Tables 5.1-5.4, the following can be observed:

- All 5D cliques contribute to the optimum set.

Table 5.2. Classification error rate for each class for the target / non-target and target type classification of the proposed algorithm (EPCM) compared to other types of classifiers using only an HH polarimetric signature as a feature.

		Co-oc matrices.				Compared classifiers			
		Soft Fusion	Hard Fusion	Opt. Set	Best 10	LNB	QNB	Fisher	KNN
Target/ non-target	Target	0.183	0.195	0.151	0.164	0.403	0.407	0.369	0.373
	Non-target	0.062	0.063	0.053	0.050	0.136	0.148	0.145	0.133
Target types	Dihedral	0.295	0.296	0.295	0.296	0.274	0.282	0.264	0.182
	Sphere	0.403	0.398	0.398	0.399	0.597	0.597	0.593	0.598
	Trihedral	0.450	0.439	0.438	0.439	0.557	0.567	0.562	0.555
	Water Jug	0.517	0.501	0.481	0.504	0.446	0.450	0.411	0.422

Table 5.3. Optimum clique sets.

Dimension	Clique numbers.	
	Target / Non-target	Target Types
2 D	1, 11, 10, 8, 2, 4.	1, 11, 10, 8, 2, 4
3 D	25, 18, 17, 24, 15, 28, 22, 27, 29.	25, 18, 17, 24, 15, 26, 27, 28, 22, 29
4 D	38, 45, 44, 50, 47, 49, 46, 41, 40	38, 45, 50, 47, 44, 49, 51, 46, 41, 40
5 D	52, 53, 54	52, 53, 54

Table 5.4. Optimum clique sets.

- Most contributions of the 2D cliques to the optimum set are diagonal cliques.
- Most of the 3 D cliques contributions to the optimum set are L-shaped cliques.
- 9-10 of the 15 cliques of the 4-variate cliques are in the optimum set.
- 9-10 of the 23 cliques of the 3-variate cliques are in the optimum set.
- 6 out of 13 cliques of the 2-variate cliques are in the optimum set.

- In the ten best arrangements, 3-variate cliques contributed only in one arrangement.
- The best 10 arrangement performance is slightly less than the optimum set performance.

5.5.2 Classification experiments using polarimetric signatures and texture features (PolTex)

5.5.2.1 Experiments description

This set of experiments follows the PolTex classification part of the scheme shown in Figure 5.1. The number of gray levels used is 12 as found in the EPCM part of the experiments. The 54 clique arrangements in the set Φ are considered.

As mentioned earlier, these experiments are performed on the cluster corresponding to the highest mean chosen after the clustering phase. Polarimetric signatures for the chosen cluster are extracted, such that each voxel has a vector of 28 polarimetric features. Texture features based on the moving window co-occurrence matrices are also computed. Each voxel will have a vector of 36 texture features. The overall feature vector will contain 64 features for each voxel. The set of all feature vectors should be checked for outliers and scaled so as to avoid unexpected behavior. This set is used as an input to the classifiers of choice. Here, we use the same set of classifiers used for comparison at the previous set of experiments. Two sets of classifiers are created, one for target /non-target classification and the other for target types classification. The voxels classified as targets are then classified as target types. The optimum set of features is sought using SFS algorithm. As the optimum set of features is different in size (number of features) for each classifier, another experiment is made using only the best 10 features for each classifier.

5.5.2.2 Experiments results

The previous steps are implemented with a 5-fold cross validation. The results for both target / non-target classification and target type classification are shown in Table 5.5. The results show that the new set of features performs much better than using only the regular HH intensity feature. Table 5.6 shows the class error rate for both target / non-target and target type classification.

Table 5.5. Classification errors for the target / non-target and target type classification using polarimetric signatures and texture features.

	LNB	LNB Opt.	LNB Best 10	QNB	QNB Opt.	QNB Best 10	Fisher	Fisher Opt.	Fisher Best 10	KNN	KNN Opt.	KNN Best 10
Target / non-target	0.085	0.084	0.089	0.116	0.087	0.088	0.085	0.083	0.092	0.028	--	0.120
Target types	0.063	0.151	0.182	0.018	0.006	0.019	0.080	0.069	0.126	0.023	--	0.133

Table 5.6. Class error rates for the target / non-target and target type classification using polarimetric signatures and texture features.

		LNB	LNB Opt.	LNB Best 10	QNB	QNB Opt.	QNB Best 10	Fisher	Fisher Opt.	Fisher Best 10	KNN	KNN Opt.	KNN Best 10
Target / non-target	Target	0.196	0.194	0.182	0.219	0.195	0.199	0.173	0.167	0.190	0.068	--	0.226
	Non-target	0.061	0.060	0.071	0.089	0.066	0.066	0.067	0.067	0.072	0.018	--	0.095
Target types	Dihedral	0.002	0.001	0.013	0.013	0.002	0.003	0.008	0.004	0.016	0.006	--	0.082
	Sphere	0.101	0.309	0.216	0.041	0.016	0.042	0.127	0.103	0.229	0.033	--	0.187
	Trihedral	0.124	0.288	0.231	0.005	0.006	0.020	0.156	0.141	0.196	0.045	--	0.192
	Water Jug	0.019	0.066	0.274	0.002	0.002	0.007	0.014	0.033	0.042	0.010	--	0.002

From Tables 5.1-5.6, we can notice that the performance of PolTex features is slightly better than EPCM in target / non-target classification while it is far better in target type classification. Among the classifiers, the KNN classifier achieves the lowest error rate (0.028) using all features for target /non-target classification, whereas Fisher's classifier is the next with an error rate of 0.084 using 47 features. QNB, which allows a minimal number of features in its optimum set for both target / non-target and target type classification, achieves the best performance for target type classification using 45 features. The performance deterioration of the QNB full set in target / non-target classification can be caused by some features that are removed on the SFS process. Last but not least, the class error rate for the target class in target / non-target classification is relatively high, except for the KNN classifier.

Table 5.7 shows the number of features in the optimum set along with the best 10 features for each classifier. For space considerations, features that didn't contribute

to the best 10 features set are omitted.

Table 5.7. Best ten features for each classifier. The best feature has the symbol $\textcircled{1}$, the next four features are labeled with the black triangle \blacktriangle and features from 6-10 are labeled with the black square \blacksquare .

		Target / non target.				Target types			
		LNB	QNB	Fisher	KNN	LNB	QNB	Fisher	KNN
Optimum Number of features		35	23	47	64	57	45	51	64
Pol. Signatures	1- T_{HH}	\blacksquare				\blacktriangle		\blacktriangle	\blacktriangle
	4- \log_{HH}	\blacksquare		\blacksquare					
	8- \log_{PHHV^*}			\blacksquare		\blacksquare	\blacktriangle	\blacktriangle	\blacksquare
	11- IR_{VVHV}	\blacktriangle	\blacksquare		\blacksquare				
	12- Re_{HHVV^*}					\blacksquare		\blacksquare	\blacktriangle
	13- Im_{HHVV^*}		\blacktriangle	\blacksquare		\blacktriangle	\blacksquare	\blacktriangle	
	18- E_{odd}	\blacktriangle	\blacktriangle	\blacktriangle	\blacktriangle			\blacksquare	
	19- E_{even}					\blacktriangle		\blacksquare	\blacktriangle
	20- med_{HH}			\blacktriangle			\blacksquare		$\textcircled{1}$
	21- max_{HH}			\blacksquare			\blacksquare		
	22- min_{HH}					\blacksquare	\blacktriangle		\blacktriangle
	23- med_{HV}	\blacktriangle							
	24- max_{HV}	\blacksquare		\blacktriangle					
	25- min_{HV}		\blacksquare			$\textcircled{1}$	$\textcircled{1}$	$\textcircled{1}$	
28- min_{VV}					\blacksquare				
Texture features from co-occurrence matrix	2 D	29- $ACOR$		\blacksquare			\blacksquare		
		31- COR		\blacksquare		\blacksquare		\blacksquare	
		34- ENT	$\textcircled{1}$		$\textcircled{1}$	\blacktriangle			
		35- HOM		$\textcircled{1}$			\blacktriangle		
	36- MP			\blacktriangle	\blacksquare				
	3 D	40- COR					\blacktriangle		
		44- HOM			\blacksquare	\blacktriangle		\blacksquare	\blacksquare
		45- MP	\blacktriangle			\blacksquare			
		46- SD	\blacksquare			\blacksquare	\blacksquare		
	4 D	49- COR				\blacktriangle			\blacksquare
5 D	56- $ACOR$		\blacktriangle		\blacktriangle			\blacksquare	
	58- COR	\blacksquare							
	59- DSM		\blacktriangle		$\textcircled{1}$		\blacksquare	\blacktriangle	\blacksquare

From Table 5.7, we can notice the following:

- Three of the texture features are counted as the *best feature* in target /non-target classification.
- The odd bounce energy feature is in the top 5 features for all classifiers in target /non-target classification.
- The \min_{HV} spatial feature is the best feature in 3 of the 4 classifiers in target type classification.
- The KNN classifier depends mostly on the texture features for target / non-target classification.
- Texture features from 4-Variate co-occurrence matrices do not contribute much to the best 10 features.
- The correlation texture feature from all k -variate co-occurrence matrices contributes to the best 10 features.

5.6 Conclusions

The problem of 3D classification of stationary targets in image-domain based Through-the-Wall Radar Imaging (TWRI) is considered. Two schemes that use the pixel-by-pixel grid and exploit the relations between pixels in 3D, or in this case voxels, have been introduced. The schemes involve classification using estimated probabilities of co-occurrence matrices and co-occurrence matrices based texture features. In addition to the new spatial intensity features, the second framework also used the polarimetric signatures. Discrimination between target / non-target classes, followed by target type classification has been demonstrated. The first framework, which uses estimated probabilities of co-occurrence matrices and fusion-based log-likelihood-ratio tests, has been proven to provide superior performance for target/non-target classification compared to state-of-the-art classifiers exploiting regular HH intensity features only. For the case of target type classification, the performance degraded. The decrease in classification performance for target type classification may be caused by the small sample size used for generating the co-occurrence matrices for each class. Small sample sizes cause sparse co-occurrence matrices, which in turn will result in poorly estimated probabilities. The second approach involves texture features, extracted by the means of a moving window co-occurrence matrix and polarimetric signatures. The polarimetric signatures include new spatial intensity features. The second set of features has been tested using four well known classifiers and has provided near perfect performance for both

target/non-target classification and target type classification. The first approach is computationally less expensive as it involves calculating the co-occurrence matrix for the whole class, followed by extracting the estimated probabilities at the testing stage. The second approach involves the computation of the polarimetric signatures, as well as the co-occurrence matrix based texture features for every voxel. EPCM method is computationally less expensive and has a run time that is less than 15% of the run time of the PolTex approach in feature extraction and training, and about 10 % less in testing. For the sake of optimization, a series of tests has been executed regarding the optimum number of gray levels and optimum set of clique arrangements. Twelve gray levels provide the lowest error rates and an optimum set of clique arrangements has been introduced. In the second framework, a series of tests to find the optimum set of features has been executed.

Chapter 6

Conclusions and Future Work

In this thesis, the problem of detecting and classifying stationary targets in Through-the-Wall Radar Imaging has been considered. Two feature-sets were introduced for TWRI imaging applications. The first depends on a combination of polarization features considering both values for each pixel or voxel by itself or considering the neighborhood of each voxel. The second feature-set makes use of the estimated probabilities or relative frequencies from the co-occurrence matrices and texture features extracted from them. A target classification framework based on polarimetric signature, the concept of super-pixels and decision trees to classify between target / non-target, target types and metal / non-metal classes in different phases is proposed. A 3D target classification framework, consisting of feature extraction, clustering and classification with exploiting polarimetric signatures and co-occurrence matrices is formulated and adapted to the application of TWRI. The features used employ 3D information and relations between voxels.

A summary and the main conclusions of the work performed in this thesis are provided in Section 6.1. Finally, Section 6.2 provides an outlook for possible future work.

6.1 Conclusions

6.1.1 Features

Various features for target detection and classification in TWRI have been investigated. Polarimetric signatures that proved to provide effective representation on Synthetic Aperture Radar (SAR) applications are implemented and a new set of spatial polarimetric signatures is introduced. The estimated probabilities of co-occurrence matrices and co-occurrence matrices based texture features showed a good performance in remote sensing and radar applications. The features based on co-occurrence matrices are extended to k-variate representations. The introduced feature-sets are tested and proved to be representative and discriminative for target detection and classification.

6.1.2 Segmentation by Classification

The problem of target segmentation and classification in the image-domain with application to Through-the-Wall Radar Imaging has been addressed. We have overcome the limitations of algorithms based on a pixel-grid, which is not a natural representation of visual scenes. Perceptually meaningful entities, obtained from a low-level grouping process, super-pixels, are considered as meaningful entities. Further, simple geometrical and statistical descriptors that have been used as features in the past are extended to polarimetric descriptors, which make use of the whole polarimetric information in radar images. A framework of polarimetric feature extraction, over-segmentation (super-pixels), clustering, and classification has been presented. An expandable sequential classifier based on random forests has been proposed to discriminate targets from clutter returns and to provide further information about the discriminated targets. The experimental results demonstrate the usefulness of the proposed methods as desired target returns are discriminable from clutter returns and a further classification to assign target types is provided. The proposed technique proved better performance than techniques that depend only on pixel intensity.

6.1.3 Multi-level 3D Classification

The problem of 3D classification of stationary targets in image-domain has been considered. Two schemes that use the pixel grid and exploit the relations between voxels in 3D were introduced. The schemes involved classification using estimated probabilities of co-occurrence matrices and co-occurrence matrices based texture features. The second framework also used the polarimetric signatures, in addition to new intensity spatial features. Discrimination between target and non-target classes followed by target type classification is demonstrated. The first framework, which used estimated probabilities of co-occurrence matrices and a fusion-based log-likelihood-ratio test, proved comparable performance for target/non-target classification to state-of-the-art classifiers. For the case of target type classification, the performance was degraded. The decrease in classification performance for target type classification may be caused by the small sample size used for generating the co-occurrence matrix for each class. Small sample sizes cause unfilled co-occurrence matrices, which in turn will result in poorly estimated probabilities. The second approach involves texture features extracted from a moving window co-occurrence matrix and polarimetric signatures. The polarimetric signatures included new spatial intensity features. The second set of features was tested using four well known

classifiers and yielded near perfect performance for both target/non-target classification and target type classification. The first approach is computationally less expensive as it involves calculating the co-occurrence matrix for the whole class followed by extracting the estimated probabilities at the testing stage. The second approach involved extracting texture features from averaged co-occurrence matrices computed over a moving window for every voxel. Also, it involved extracting polarimetric signatures for the involved voxels. EPCM method is computationally less expensive than the PolTex approach. For the sake of optimization, a series of tests were executed for both of the frameworks. In the first framework, test experiments for the optimum number of gray levels for image quantization before computing the co-occurrence matrices were carried out. Twelve gray levels provided the lowest error rates. Also, experiments to choose the optimum set of clique arrangement were executed. In the second framework, a series of tests to find the optimum set of features were executed, and a table that points out the best 10 features associated with each classifier is introduced. The optimum set of clique arrangements or features or even the optimum number of gray levels may differ when changing the training set or classifier. We suggest to optimize the experiment parameters according to the test environment. The algorithms introduced in this section can be adapted to other image processing applications, especially applications with non-optical nature as in medical imaging.

6.2 Future Work

6.2.1 Super-voxels

In this thesis, we have exploited super-pixels as homogeneous meaningful entities. The extension of the super-pixels approach to super-voxels can benefit from the correlated information between pixels on 3D (voxels) and may better enhance the performance of detection and classification.

6.2.2 Clutter and noise reduction

In TWRI, the importance of clutter and noise reduction techniques has to be explored [55]. Clutter reduction techniques are classified among others as statistical signal processing [156], classical filtering [157, 158] and non linear signal processing based on neural networks [159]. Automatic clutter reduction based on statistical

and multilayer perceptrons is described in [160]. Clutter reduction based on statistical signal processing techniques such as PCA and ICA [161] and SVD is considered in [55] to remove the clutter. Clutter reduction as a pre-processing step will enhance the performance of detection and classification algorithms. Thus, a joint detection and clutter reduction scheme could help to improve the detection result.

6.2.3 Features

Little work has been done in feature extraction for TWRI images. First attempts in this field include the work by Debes et al. [84,85] and the work by the author of this thesis [87]. The use of texture features proved to be efficient for target classification as we tested in this thesis. Other types of texture features [162] may add information that contribute to target classification. Semantic and relational features [163] can provide some context to the scene under test exploiting other features like shape, texture, and material. Shape features, such as superquadrics and its extension using global deformations [164], sets of concatenated superquadrics [165] or free-form deformations [166–168] can provide discriminative information between 3D objects. As we experienced here, a lot of features can be used in target detection and classification, but increasing the number of features usually comes with a decrease in the performance which is known as the curse of dimensionality [169]. A feature selection algorithm is desirable to have a discriminant subset of features without losing performance.

6.2.4 Classification

In this thesis, although we used unsupervised learning approaches like cluster analysis, we have restricted ourselves to supervised classification approaches. Unsupervised learning may be more appropriate in some situations than supervised learning, as in many cases the parameters needed for supervised learning may depend on knowledge that are not available.

Further, we used simple linear and non-linear classifiers as K -nearest neighbor, Quadratic Normal Bayes, and Random forests. Using Support Vector Machines [170,171] and Neural Networks [172] might be more appropriate for more complex scenarios. In [84,85], methods for feature compensation that are invariant to translation and system resolution are proposed. Rotation-invariance is not considered so far. Training the classifier with data collected from different vantage points can slightly compensate the effect. An artificially rotated data set can also be used for

training, but rotated object may belong to different class or sub-class of the original one in order to decrease misclassification rate. Model-based approaches that model its parameters as a function of the viewing angle can be considered.

6.2.5 Wall Removal

Having empty room measurements available is an ideal case which may be inappropriate in practical scenarios. First studies on the effect of wall removal on detection are published in [76] and [173]. How wall removal techniques affect detection and classification results is still an unanswered question [11].

List of Acronyms

2D	Two dimensional
3D	Three dimensional
5d MRF	5-variate Markov Random Field
ACOR	Auto Correlation
AUC	Area Under the Curve
CART	Classification And Regression Trees
CON	Contrast
COR	Correlation
DSBF	Delay-and-sum beam former
DSM	Dissimilarity
EM	Electromagnetic
ENT	Entropy
EPCM	Estimated probabilities of co-occurrence matrices
FLS	Fishers Least Square Linear Classifier
GLCM	Gray level co-occurrence matrix
HH	Horizontal return, horizontal transmit
HOM	Homogeneity
HV	Horizontal return, vertical transmit
ICA	Independent Component Analysis
ICM	Iterated Conditional Modes
KNN	K-Nearest Neighbor Classifier [119, 139]
LBN	Linear Bayes Normal Classifier [119,138] ,
LSM	Levelset Method
MIMO	Multiple Input Multiple Output

MP	Maximum Probability
MRF	Markov Random Fields
MRI	Magnetic Resonance Imaging
MSE	Mean Square Error
MVDR	Minimum variance distortion-less response
OOB	Out-Of-Bag
PCA	Principal Component Analysis
pdf	probability density function
PoITEX	POlarimetric signatures and TEXture features
PSF	Point Spread Function
QBN	Quadratic Bayes Normal Classifier
RCS	Radar Cross Section
RF	Radio Frequency
ROC	Receiver Operating Characteristic
SAR	Synthetic Aperture Radar
SD	Standard Deviation
SNR	Signal-to-Noise Ratio
SFS	Sequential Forward Selection
SVD	Singular Value Decomposition
TWRI	Through-the-Wall Radar Imaging
VV	Vertical return, vertical transmit

List of Symbols

\mathbf{AP}_{RI}	Vector of real and imaginary parts of normalized product of different polarimetric channels
\mathbf{C}_1	Polarimetric covariance matrix
$\mathbf{C}(i, j)$	i th row and j th column of the co-occurrence matrix
D	number of features
ξ_s	set of pixels belonging to a super pixel s
E_{even}	Even bounce energy
E_{odd}	Odd bounce energy
f_m	m -th step frequency
\mathbf{F}_q	feature vector of the q th pixel
\mathbf{FS}_s	feature vector of the super pixel s
$\mathbf{FP}(q)$	Polarimetric feature vector at voxel q
G	number of gray levels in an image or a co-occurrence matrix
$g(x, y, z)$	Gray level function
i_k	Gray level of the voxel k on the clique arrangement tested
Im_{HHVV^*}	Imaginary part of normalized product between two complex polarized components HH and VV
\mathbf{IP}_{\log}	Logarithmic features vector
\mathbf{IR}	Normalized intensity ratio vector
IR_{HHVV}	Intensity ratio between two complex polarized components HH and VV
IR_{HHHV}	Intensity ratio between two complex polarized components HH and HV
IR_{VVHV}	Intensity ratio between two complex polarized components VV and HV
\mathcal{L}	Lattice
\hat{L}	estimated log-likelihood ratio
L_{ϕ_r}	log-likelihood ratio per clique ϕ_r
Λ	Displacement vector
λ_x^k	Displacement in x direction of the element k on a clique from the center voxel
\log_{HH}	Logarithmic intensity of the HH polarized component
\log_{HHVV^*}	Magnitude of the HH - VV correlation function
\log_{PHHV^*}	HH - VV phase difference
μ_1	Mean of the marginal probability p_1 for 2-variate co-occurrence matrix

μ_m	Mean of the marginal probability p_m for k -variate co-occurrence matrix
\mathcal{N}	Neighborhood system
\hat{O}_b	predicted class label by the b th tree
O_b	actual class label
Ω_b	set of classes
ω_{mb}	class with index m on the set Ω
$\hat{p}_{i_1, \dots, i_k}$	Estimated k -variate pdf
$P(v_i v_{\mathcal{N}_i})$	Probability of an element conditioned to its neighbors
P_{F_r}	Probability of false alarm per clique ϕ_r
P_{M_r}	Probability of miss detection per clique ϕ_r
Φ	Set of cliques
ϕ_r	clique with index r in the set Φ
Ψ	Normalized phase difference
q	voxel in the region of interest
R	Number of cliques / random variables
$R_k(u', v')$	Distance from the k -th antenna
Re_{HHVV^*}	Real part of normalized product between two complex polarized components HH and VV
Re_{HHHV^*}	Real part of normalized product between two complex polarized components HH and HV
Re_{HVVV^*}	Real part of normalized product between two complex polarized components HV and VV
S	number of super-pixels
S_{HV}	Complex amplitude of the H - (horizontal) polarized return given that the transmitted signal is V - (vertically) polarized
σ_p	reflectivity of discrete targets
SP	Spatial intensity features vector
T	Thresholded intensity features vector
τ_{pn}	round trip delay between p -th target and n -th receiver
τ_{qn}	delay compensation for n -th receiver steering beam at position q
u_r	decision in fusion
w_m	weighting factor
χ	Set of random variables

Bibliography

- [1] M. G. Amin, ed., *Through-the-Wall Radar Imaging*. CRC Press, 2011.
- [2] F. Ahmad and M. Amin, “Multi-location wideband synthetic aperture imaging for urban sensing applications,” *Journal of the Franklin Institute*, vol. 345, pp. 618–639, Sep 2008.
- [3] D. Ferris and N. Currie, “A survey of current technologies for through-the-wall surveillance TWS,” in *Proceedings of the SPIE Conference on Sensors, C31 Information, and Training Technologies for Law Enforcement*, vol. vol. 3577, (Boston, MA, USA), pp. 62–72, 1998.
- [4] M. Amin and K. Sarabandi, “Special issue on remote sensing of building interior,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 1267–1268, May 2009.
- [5] M. Amin (Guest Editor), “Special issue: Advances in indoor radar imaging,” *Journal of the Franklin Institute*, vol. 345, no. 6, 2008.
- [6] E. Baranoski and F. A. S. Chairs), “Special session on through-the-wall radar imaging,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008.
- [7] C. Debes, M. Amin, and A. Zoubir, “Target detection in single- and multiple-view through-the-wall radar imaging,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 1349–1361, May 2009.
- [8] E. Baranoski, “Through-wall imaging: Historical perspective and future directions,” *Journal of the Franklin Institute*, vol. 345, pp. 556–569, Sep 2008.
- [9] R. Solimene, F. Soldovieri, G. Prisco, and R. Pierri, “Three-dimensional through-wall imaging under ambiguous wall parameters,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 1310–1317, May 2009.
- [10] T. Dogaru and C. Le, “Through-the-wall small weapon detection based on polarimetric radar techniques,” tech. rep., Army Research Laboratory, Adelphi, MD, 2009.
- [11] C. Debes, *Advances in Detection and Classification for Through-the-Wall Radar Imaging*. PhD thesis, Technische Universit Darmstadt, 8 2010.
- [12] P. Setlur, F. Ahmad, and M. Amin, “Maximum likelihood and suboptimal schemes for micro-doppler estimation using carrier diverse doppler radars,” *IET Signal Processing*, vol. 5, pp. 194–208, April 2011.
- [13] R. Solimene, R. Di Napoli, F. Soldovieri, and R. Pierri, “TWI for an unknown symmetric lossless wall,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 8, pp. 2876–2886, 2011.

-
- [14] M. Thiel and K. Sarabandi, "Ultra-wideband multi-static scattering analysis of human movement within buildings for the purpose of stand-off detection and localization," *IEEE Trans. on Antennas and Propagation*, vol. 59, no. 4, 2011.
- [15] S. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, pp. 498–518, Mar 1996.
- [16] P. Yu, A. K. Qin, and D. A. Clausi, "Unsupervised polarimetric SAR image segmentation and classification using region growing with edge penalty," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, no. 99, pp. 1–16, 2011.
- [17] G. Zhou, Y. Cui, Y. Chen, J. Yang, H. Rashvand, and Y. Yamaguchi, "Linear feature detection in polarimetric SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 1453–1463, April 2011.
- [18] R. Paladini, M. Martorella, and F. Berizzi, "Classification of man-made targets via invariant coherency-matrix eigenvector decomposition of polarimetric SAR/ISAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 3022–3034, Aug. 2011.
- [19] Jong-Sen Lee, M. Grunes, and E. Pottier, "Quantitative comparison of classification capability: fully polarimetric versus dual and single-polarization SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 11, pp. 2343–2351, 2001.
- [20] Jong-Sen Lee, K. Hoppel, S. Mango, and A. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 5, pp. 1017–1028, 1994.
- [21] D. Hoekman and M. Vissers, "A new polarimetric classification approach evaluated for agricultural crops," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 2881–2889, Dec 2003.
- [22] M. Petrou and P. Sevilla, *Image processing: dealing with texture*. John Wiley & Sons Inc., 2006.
- [23] S. Shang, X. Kong, and X. You, "Copy paper brand source identification using commodity scanners," *Information Technology Journal*, vol. 10, 2011.
- [24] X. Sheng, P. Qi-Cong, H. Shan, and U. Center, "An invariant descriptor design method based on MSER," *Information Technology Journal*, vol. 9, no. 7, pp. 1345–1352, 2010.
- [25] B. Mingde, S. Zhigang, and L. Yesong, "Textural fabric defect detection using adaptive quantized gray-level co-occurrence matrix and support vector description data," *Information Technology Journal*, vol. 11, pp. 673–685, 2012.

- [26] R. Tao, J. Chen, B. Chen, and C. Liu, "GLCM and fuzzy clustering for ocean features classification," in *International Conference on Machine Vision and Human-Machine Interface (MVHI)*, pp. 538–540, IEEE, 2010.
- [27] A. Muqaibel, A. Safaai-Jazi, A. Bayram, M. Attiya, and S. Riad, "Ultrawide-band through-the-wall propagation," *IEE Proceedings Microwave, Antennas and Propagation*, vol. 152, pp. 581–588, 2005.
- [28] J. Ren, S. Chai, and W. Chen, "Modeling the UWB through-the-wall signal pathloss measurements and time domain ray tracing method," in *2011 IEEE International Conference on Microwave Technology Computational Electromagnetics (ICMTCE)*, pp. 334–337, May 2011.
- [29] I. Akduman, L. Crocco, and F. Soldovieri, "Experimental validation of a simple system for through-the-wall inverse scattering," *Geoscience and Remote Sensing Letters, IEEE*, vol. 8, pp. 258–262, March 2011.
- [30] W. Zheng, Z. Zhao, and Z. Nie, "Application of trm in the UWB through wall radar," *Progress In Electromagnetics Research*, vol. 87, pp. 279–296, 2008.
- [31] F. Soldovieri and R. Solimene, "Through-wall imaging via a linear inverse scattering algorithm," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, pp. 513–517, 2007.
- [32] R. Narayanan, "Through-wall radar imaging using UWB noise waveforms," *Journal of the Franklin Institute*, vol. 345, pp. 659–678, Sept 2008.
- [33] L. Huang and Y. Lu, "Through-the-wall tomographic imaging using chirp signals," in *2011 IEEE International Symposium on Antennas and Propagation (APSURSI)*, pp. 2091–2094, July 2011.
- [34] E. Ertin, R. Moses, and R. Burkholder, "Autofocus for coherent through-the-wall imaging with multiple antenna arrays," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on*, pp. 69–72, Dec. 2011.
- [35] P.-H. Chen, M. C. Shastry, C.-P. Lai, and R. M. Narayanan, "A portable real-time digital noise radar system for through-the-wall imaging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, no. 99, pp. 1–12, 2012.
- [36] T. Kumar and P. Kumar, "UWB impulse radar for three dimensional through-the-wall radar imaging," in *Microwaves, Communications, Antennas and Electronics Systems (COMCAS), 2011 IEEE International Conference on*, pp. 1–5, Nov. 2011.
- [37] L.-P. Song, C. Yu, and Q. H. Liu, "Through-wall imaging (TWI) by radar: 2-D tomographic results and analyses," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 12, pp. 2793–2798, 2005.
- [38] R. Solimene, F. Soldovieri, G. Prisco, and R. Pierri, "Three-dimensional microwave tomography by a 2-d slice-based reconstruction algorithm," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 4, pp. 556–560, 2007.

- [39] F. Ahmad and M. Amin, "Noncoherent approach to through-the-wall radar localization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, pp. 1405–1419, Oct 2006.
- [40] M. Dehmollaian and K. Sarabandi, "Refocusing through building walls using synthetic aperture radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 6, pp. 1589–1599, 2008.
- [41] Y.-S. Yoon and M. G. Amin, "High-resolution through-the-wall radar imaging using beamspace music," *IEEE Transactions on Antennas and Propagation*, vol. 56, pp. 1763–1774, Jun 2008.
- [42] M. Aamna, S. Ammar, T. Rameez, S. Shabeeb, N. Rao, and I. Safwat, "2d beamforming for through-the-wall microwave imaging applications," in *International Conference on Information and Emerging Technologies (ICIET)*, pp. 1–6, June 2010.
- [43] W. Zhang, A. Hoorfar, and C. Thajudeen, "Building layout and interior target imaging with SAR using an efficient beamformer," in *2011 IEEE International Symposium on Antennas and Propagation (APSURSI)*, pp. 2087–2090, July 2011.
- [44] W. Zhang, C. Thajudeen, and A. Hoorfar, "Polarimetric through-the-wall imaging," *2010 URSI International Symposium on Electromagnetic Theory (EMTS)*, pp. 471 – 474, 2010.
- [45] B. Boudamouz, P. Millot, and C. Pichot, "Through the wall radar imaging with mimo beamforming processing," in *Microwaves, Radar and Remote Sensing Symposium (MRRS)*, pp. 251–254, Aug. 2011.
- [46] Y.-S. Yoon, M. Amin, and F. Ahmad, "MVDR beamforming for through-the-wall radar imaging," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, pp. 347–366, Jan. 2011.
- [47] S. Gauthier, E. Hung, and W. Chamma, "Surveillance through concrete walls," in *Proceedings of SPIE 5403*, pp. 597–608, 2004.
- [48] F. Ahmad, Y. Zhang, and M. Amin, "Three-dimensional wideband beamforming for imaging through a single wall," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 176–179, 2008.
- [49] L. Li, W. Zhang, and F. Li, "A novel autofocusing approach for real-time through-wall imaging under unknown wall characteristics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 1, pp. 423–431, 2010.
- [50] G. Wang and M. Amin, "Imaging through unknown walls using different stand-off distances," *IEEE Transactions on Signal Processing*, vol. 54, pp. 4015–4025, Oct 2006.
- [51] F. Tivive, A. Bouzerdoum, and M. Amin, "An SVD-based approach for mitigating wall reflections in through-the-wall radar imaging," in *IEEE Radar Conference (RADAR)*, pp. 519–524, May 2011.

- [52] M. Dehmollaian, M. Thiel, and K. Sarabandi, "Through-the-wall imaging using differential SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1289–1296, 2009.
- [53] R. Chandra, A. Gaikwad, D. Singh, and M. Nigam, "An approach to remove the clutter and detect the target for ultra-wideband through-wall imaging," *Journal of Geophysics and Engineering*, vol. 5, p. 412, 2008.
- [54] A. A. Mostafa and A. Zoubir, "3d target detection in through-the-wall radar imaging," in *Proceedings of SPIE, Signal Processing, Sensor Fusion, and Target Recognition XIX*, vol. 7697, p. 76971F, 2010.
- [55] P. Verma, A. Gaikwad, D. Singh, and M. Nigam, "Analysis of clutter reduction techniques for through wall imaging in UWB range," *Progress In Electromagnetics Research*, vol. 17, pp. 29–48, 2009.
- [56] M. Leigsnering and A. M. Zoubir, "Fast wideband near-field imaging using the non-equispaced FFT with application to through-wall radar," in *19th European Signal Processing Conference (EUSIPCO)*, 8 2011.
- [57] M. Nikolic, A. Nehorai, and A. Djordjevic, "Sparse through-the-wall imaging," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on*, pp. 77–80, Dec. 2011.
- [58] Y.-S. Yoon and M. Amin, "Through-the-wall radar imaging using compressive sensing along temporal frequency domain," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 2806–2809, March 2010.
- [59] Y.-S. Yoon and M. G. Amin, "Imaging of behind the wall targets using wide-band beamforming with compressive sensing," in *IEEE International Workshop on Statistical Signal Processing*, pp. 93–96, 2009.
- [60] F. Ahmad, M. Amin, and G. Mandapati, "Autofocusing of through-the-wall radar imagery under unknown wall characteristics," *IEEE Transactions on Image Processing*, vol. 16, no. 7, pp. 1785–1795, 2007.
- [61] F. Ahmad, T. Tchavov, S. Kassam, and M. Amin, "Image quality measures for wall parameter estimation," in *Proceedings of the IEEE International Symposium on Signal Processing and Information Technology*, pp. 676–679, 2003.
- [62] Z. Rosenbaum and B. G. Mobasserri, "Interpretation of through-the-wall radar imagery by probabilistic volume model building," in *Proceedings of SPIE, the International Society for Optical Engineering*, vol. 6943, pp. 69430T–10, SPIE Defense, Security And Sensing, 2008.
- [63] S. Hantscher, A. Reizenzahn, and C. Diskus, "Through-wall imaging with a 3-D UWB SAR algorithm," *IEEE Signal Processing Letters*, vol. 15, pp. 269–272, 2008.

- [64] B. Mobasser, G. Smith, and I. Estephan, "A target alignment algorithm for through-the-wall radar imagery classification," in *Proceedings of the IEEE International Radar Conference*, pp. 756–761, 2010.
- [65] N. Maaref, P. Millot, P. Pichot, and O. Picon, "A study of UWB FM-CW radar for the detection of human beings in motion inside a building," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1297–1300, 2009.
- [66] Y. Kim and H. Ling, "Human activity classification based on micro-doppler signatures using a support vector machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1328–1337, 2009.
- [67] A. Lin and H. Ling, "Through-wall measurements of a doppler and direction-of-arrival DDOA radar for tracking indoor movers," in *Proceedings of the IEEE Antennas and Propagation Society*, vol. 3B, pp. 322–325, 2005.
- [68] P. Setlur, F. Ahmad, M. Amin, and P. Zeman, "Experiments on through-the-wall motion detection and ranging," in *Proceedings of the SPIE*, 2007.
- [69] A. Buonanno, M. D'Urso, M. Felaco, G. Prisco, L. Angrisani, M. Ascione, and R. Moriello, "Through-the-wall tracking of moving targets: Tests on experimental data," in *IEEE Radar Conference (RADAR)*, pp. 846–849, May 2011.
- [70] K. Chetty, G. Smith, and K. Woodbridge, "Through-the-wall sensing of personnel using passive bistatic wifi radar at standoff distances," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 1218–1226, April 2012.
- [71] A. Martone, C. Le, and K. Ranney, "Coherent and non-coherent change detection for through the wall sensing of moving targets," in *IEEE Radar Conference*, pp. 256–261, May 2011.
- [72] C. Debes, M. Amin, and A. Zoubir, "Target detection in multiple-viewing through-the-wall radar imaging," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pp. 173–176, 2008.
- [73] C. Debes, J. Riedler, M. G. Amin, and A. M. Zoubir, "Iterative target detection approach for through-the-wall radar imaging," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3061–3064, Apr 2009.
- [74] C. Debes, A. Zoubir, and M. Amin, "Optimal decision fusion in through-the-wall radar imaging," in *Proceedings of the IEEE International Workshop on Statistical Signal Processing*, pp. 761–764, 2009.
- [75] C. Debes, C. Weiss, M. Amin, and A. Zoubir, "Distributed target detection in through-the-wall radar imaging using the bootstrap," in *Proceedings of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 3530–3533, 2010.

- [76] C. Debes, J. Riedler, M. G. Amin, and A. M. Zoubir, "Adaptive target detection with application to through-the-wall radar imaging," *IEEE Transactions on Signal Processing*, vol. 58, pp. 5572–5583, Nov 2010.
- [77] C. Debes and A. Zoubir, *Through-the-Wall Radar Imaging*, ch. Detection Approaches in Through Wall Radar Imaging. CRC Press, 2010.
- [78] C. Debes and A. Zoubir, "Image-domain based target detection under model uncertainties in through-the-wall radar imaging," in *Proceedings of the American Electromagnetics Conference*, 2010.
- [79] C. Debes, A. M. Zoubir, and M. G. Amin, "Detection of geometrically known targets in through-the-wall radar imaging," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3624–3627, May 2011.
- [80] C. Debes, A. M. Zoubir, and M. G. Amin, "Enhanced detection using target polarization signatures in through-the-wall radar imaging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 1968–1979, May 2012.
- [81] F. Ahmad, "Multi-location wideband through-the-wall beamforming," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008.
- [82] C. Seng, A. Bouzerdoum, S. Phung, and M. Amin, "A two-stage image fusion method for enhanced through-the-wall radar target detection," in *IEEE Radar Conference (RADAR)*, pp. 643–647, May 2011.
- [83] G. Smith and B. Mobasseri, "Resolution independent radar target recognition," in *IEEE Radar Conference (RADAR)*, pp. 905–910, May 2011.
- [84] C. Debes, J. Hahn, M. Amin, and A. Zoubir, "Feature extraction in through-the-wall radar imaging," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3562–3565, 2010.
- [85] C. Debes, J. Hahn, A. Zoubir, and M. Amin, "Target discrimination and classification in through-the-wall radar imaging," *IEEE Transactions on Signal Processing*, vol. 59, pp. 4664–4676, Oct. 2011.
- [86] B. G. Mobasseri and Z. Rosenbaum, "3D classification of through-the-wall radar images using statistical object models," in *Proceedings of the IEEE Workshop on Image Analysis and Interpretation*, pp. 149–152, 2008.
- [87] A. A. Mostafa, C. Debes, and A. M. Zoubir, "Segmentation by classification for through-the-wall radar imaging using polarization signatures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 3425–3439, Sept 2012.
- [88] G. Alli and D. DiFilippo, *Through-the-Wall Radar Imaging*, ch. Beamforming for Through-the-Wall Radar Imaging. CRC Press, 2010.

- [89] F. Ahmad and M. Amin, "High-resolution imaging using capon beamformers for urban sensing applications," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 985–988, 2007.
- [90] M. Leigsnering, C. Debes, and A. M. Zoubir, "Compressive sensing in through-the-wall radar imaging," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4008–4011, May 2011.
- [91] D. Massonnet and J.-C. Souyris, *Synthetic Aperture Radar Imaging*. EFPL Press, May 2008.
- [92] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. SciTech, 2004.
- [93] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging, From Basics to Applications*, vol. 143 of *Optical Science and Engineering*. CRC Press, Feb 2009.
- [94] H. Zebker and J. Van Zyl, "Imaging radar polarimetry: a review," *Proceedings of the IEEE*, vol. 79, no. 11, pp. 1583–1606, 1991.
- [95] R. Vaccaro, P. Smits, and S. Dellepiane, "Exploiting spatial correlation features for SAR image analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 1212–1223, May 2000.
- [96] R. Vaccaro and S. Dellepiane, "Features estimation under sample dependence hypothesis," tech. rep., Department of Biophysics Electrical Engineering, University of Genoa, Genoa, Italy, 1998.
- [97] R. Vaccaro and S. Dellepiane, "Spatial correlation features for SAR images in a small sample size context Image Analysis and Processing," vol. 1311 of *Lecture Notes in Computer Science*, ch. 70, pp. 560–567, Berlin, Heidelberg: Springer Berlin / Heidelberg, 1997.
- [98] F. Ulaby, F. Kouyate, B. Brisco, and T. Williams, "Textural information in SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-24, no. 2, pp. 235–245, 1986.
- [99] W. Schwartz and H. Pedrini, "Texture classification based on spatial dependence features using co-occurrence matrices and markov random fields," in *International Conference on Image Processing (ICIP'04.)*, vol. 1, pp. 239–242, IEEE, 2004.
- [100] L.-K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, pp. 780–795, March 1999.
- [101] J. Lai and F. Tsai, "Three dimensional texture computation of gray level co-occurrence tensor in hyperspectral image cubes," in *Asian Conference on Remote Sensing*, Asian Association of Remote Sensing, 2008. Section TS16.2.

- [102] E. Rignot, R. Chellappa, and P. Dubois, "Unsupervised segmentation of polarimetric SAR data using the covariance matrix," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, pp. 697–705, Jul 1992.
- [103] D. E. Kreithen, S. S. Halversen, and G. J. Owirka, "Discriminating targets from clutter," *Lincoln Lab. J.*, vol. 6, no. 1, pp. 25–52, 1993.
- [104] L. Novak, S. Halversen, G. Owirka, and M. Hielt, "Effects of polarization and resolution on the performance of a SAR automatic target recognition system," *The Lincoln Laboratory Journal*, vol. 8, no. 1, pp. 49–68, 1995.
- [105] L. Novak, S. Halversen, G. Owirka, and M. Hielt, "Effects of polarization and resolution on SAR atr," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, pp. 102–116, Jan. 1997.
- [106] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.
- [107] W. Lv, Q. Yu, and W. Yu, "Water extraction in SAR images using GLCM and support vector machine," in *IEEE 10th International Conference on Signal Processing (ICSP)*, pp. 740–743, IEEE, 2010.
- [108] F. Tsai, C. Chang, J. Rau, T. Lin, and G. Liu, "3d computation of gray level co-occurrence in hyperspectral image cubes," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 429–440, Springer, 2007.
- [109] W. Chen, R. Huang, and L. Hsieh, "Iris recognition using 3d co-occurrence matrix," *Advances in Biometrics*, pp. 1122–1131, 2009.
- [110] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society. Series B*, vol. 36, no. 2, pp. 192–236, 1974.
- [111] R. Kashyap and R. Chellappa, "Estimation and choice of neighbors in spatial-interaction models of images," *IEEE Transactions on Information Theory*, vol. 29, pp. 60–72, Jan 1983.
- [112] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, pp. 721–741, Nov. 1984.
- [113] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-5, pp. 25–39, Jan. 1983.
- [114] D. Clausi and Y. Zhao, "An advanced computational method to determine co-occurrence probability texture features," in *IEEE International Geoscience and Remote Sensing Symposium, (IGARSS)*, vol. 4, pp. 2453–2455 vol.4, June 2002.

- [115] J. Mridula, K. Kumar, and D. Patra, "Combining GLCM features and markov random field model for colour textured image segmentation," in *International Conference on Devices and Communications (ICDeCom)*, pp. 1–5, IEEE, 2011.
- [116] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, October 2001.
- [117] X. Ren and J. Malik, "Learning a classification model for segmentation," in *Proceedings of the IEEE Ninth International Conference on Computer Vision*, pp. 10–17 vol.1, 2003.
- [118] B. Fulkerson, A. Vedaldi, and S. Soatto, "Class segmentation and object localization with superpixel neighborhoods," in *IEEE 12th International Conference on Computer Vision*, pp. 670–677, IEEE, Sep 2009.
- [119] D. Thompson, L. Mandrake, M. Gilmore, and R. Castañando, "Superpixel endmember detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, pp. 4023–4033, Nov. 2010.
- [120] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Wadsworth, 1984.
- [121] D. L. Verbyla, "Classification trees: a new discrimination tool," *Canadian Journal of Forest Research*, vol. 17, pp. 1150–1152, Sep 1987.
- [122] L. Clark and D. Pregibon, "Tree-based models. statistical models in S," *Wadsworth, Pacific Grove*, 1992.
- [123] V. Svetnik, A. Liaw, C. Tong, J. Culberson, R. Sheridan, and B. Feuston, "Random forest: a classification and regression tool for compound classification and qsar modeling," *Journal of Chemical Information and Computer Sciences*, vol. 43, no. 6, pp. 1947–1958, 2003.
- [124] A. Vedaldi and S. Soatto, *Quick Shift and Kernel Methods for Mode Seeking*, vol. 5305 of *Lecture Notes in Computer Science*, ch. 52, pp. 705–718. Springer Berlin Heidelberg, 2008.
- [125] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms." <http://www.vlfeat.org/>, 2008.
- [126] R. Duda, P. Hart, and D. Stork, *Pattern classification*, vol. 2. John Wiley and Sons, 2001.
- [127] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer Verlag, 2009.
- [128] C. Gini, "Variability and mutability," *contribution to the study of statistical distributions and relations, Studi Economico-Giuridici della R. Universita de Cagliari.*, 1912.
- [129] R. Timofeev, "Classification and regression trees (CART) theory and applications." <http://edoc.hu-berlin.de/docviews/abstract.php?id=26951>, 2004. [Online: Stand 2011-05-23T08:37:58Z].

- [130] T. Dietterich, *Ensemble Learning, Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, 2002.
- [131] A. M. Zoubir and D. R. Iskander, *Bootstrap Techniques for Signal Processing*. Cambridge University Press, 2004.
- [132] L. Breiman, “Out-of-bag estimation,” *Technical Report, Department of Statistics, UC Berkeley*, 1996.
- [133] T. Fawcett, “An introduction to ROC analysis,” *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [134] C.-P. Lai and R. M. Narayanan, “Ultrawideband random noise radar design for through-wall surveillance,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, pp. 1716–1730, Oct 2010.
- [135] H. Estephan, M. G. Amin, and K. M. Yemelyanov, “Optimal waveform design for improved indoor target detection in sensing through-the-wall applications,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, pp. 2930–2941, Jul 2010.
- [136] F. Ahmad and M. G. Amin, “Matched-illumination waveform design for a multistatic through-the-wall radar system,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, pp. 177–186, Feb 2010.
- [137] F. Ahmad and M. Amin, “Through-the-wall polarimetric imaging,” *Proceedings of SPIE, Algorithms for Synthetic Aperture Radar Imagery XV*, vol. 6970, pp. 69700N–1–69700N–10, 2008.
- [138] Y. Wang, Y. Yang, and A. Fathy, “Experimental assessment of the cross coupling and polarization effects on ultra-wide band see-through-wall imaging reconstruction,” *Microwave Symposium Digest, IEEE MTT-S Int.*, pp. 9–12, 2009.
- [139] K. M. Yemelyanov, N. Engheta, A. Hoorfar, and J. A. McVay, “Adaptive polarization contrast techniques for through-wall microwave imaging applications,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 1362–1374, May 2009.
- [140] J. MacQueen *et al.*, “Some methods for classification and analysis of multivariate observations,” in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, pp. 281–297, California, USA, 1967.
- [141] P. Setlur, M. Amin, and F. Ahmad, “Multipath model and exploitation in through-the-wall and urban radar sensing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 4021–4034, 2011.
- [142] E. Forgy, “Cluster analysis of multivariate data: efficiency versus interpretability of classifications,” *Biometrics*, vol. 21, pp. 768–769, 1965.
- [143] R. Xu, D. Wunsch, *et al.*, “Survey of clustering algorithms,” *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 645–678, 2005.

- [144] X. Zhang and P. Rosin, "Superellipse fitting to partial data," *Pattern Recognition*, vol. 36, no. 3, pp. 743–752, 2003.
- [145] P. Varshney and C. Burrus, *Distributed detection and data fusion*. Springer Verlag, 1997.
- [146] M. Kudo and J. Sklansky, "Comparison of algorithms that select features for pattern classifiers," *Pattern Recognition*, vol. 33, no. 1, pp. 25–41, 2000.
- [147] S. Knerr, L. Personnaz, and G. Dreyfus, "Single-layer learning revisited: A stepwise procedure for building and training a neural network," in *Neurocomputing: Algorithms, Architectures and Applications* (F. Fogelman Soulié and J. Héroult, eds.), vol. F68 of *NATO ASI Series*, pp. 41–50, Springer-Verlag, 1990.
- [148] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, pp. 415–425, March 2002.
- [149] K. Duan and S. Keerthi, "Which is the best multiclass SVM method? an empirical study," *Multiple Classifier Systems*, vol. LNCS 3541, pp. 278–285, 2005.
- [150] K. Fukunaga, *Introduction to statistical pattern recognition*. Academic Press, 1990.
- [151] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, pp. 21–27, Jan. 1967.
- [152] R. A. FISHER, "The use of multiple measurements in taxonomic problems," *Annals of Human Genetics*, vol. 7, no. 2, pp. 179–188, 1936.
- [153] S. Raudys and R. P. W. Duin, "Expected classification error of the fisher linear classifier with pseudo-inverse covariance matrix," *Pattern Recogn. Lett.*, vol. 19, pp. 385–392, Apr. 1998.
- [154] P. Forero, V. Kekatos, and G. Giannakis, "Robust clustering using outlier-sparsity regularization," *IEEE Transactions on Signal Processing*, 2012. Early access.
- [155] H.-G. Li, G.-Q. Wu, X.-G. Hu, J. Zhang, L. Li, and X. Wu, "K-means clustering with bagging and mapreduce," in *44th Hawaii International Conference on System Sciences (HICSS)*, pp. 1–8, Jan. 2011.
- [156] X. Xu, E. Miller, C. Rappaport, and G. Sower, "Statistical method to detect subsurface objects using array ground-penetrating radar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 4, pp. 963–976, 2002.
- [157] A. Van Der Merwe and I. Gupta, "A novel signal processing technique for clutter reduction in GPR measurements of small, shallow land mines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 6, pp. 2627–2637, 2000.

- [158] A. Zoubir, I. Chant, C. Brown, B. Barkat, and C. Abeynayake, "Signal processing techniques for landmine detection using impulse ground penetrating radar," *IEEE Sensors Journal*, vol. 2, no. 1, pp. 41–51, 2002.
- [159] R. Vicen-Bueno, R. Carrasco-Álvarez, M. Rosa-Zurera, and J. Nieto-Borge, "Sea clutter reduction and target enhancement by neural networks in a marine radar system," *Sensors*, vol. 9, no. 3, pp. 1913–1936, 2009.
- [160] N. Xie, H. Leung, and H. Chan, "A multiple-model prediction approach for sea clutter modeling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1491–1502, 2003.
- [161] F. Abujarad and A. Omar, "GPR data processing using the component-separation methods pca and ica," in *Imagining Systems and Techniques, 2006. IST 2006. Proceedings of the 2006 IEEE International Workshop on [Imagining read Imaging]*, pp. 60–64, IEEE, 2006.
- [162] T. Randen and J. Husoy, "Filtering for texture classification: A comparative study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 291–310, 1999.
- [163] J. Vogel and B. Schiele, "Natural scene retrieval based on a semantic modeling step," *Image and Video Retrieval*, pp. 1950–1950, 2004.
- [164] F. Solina and R. Bajcsy, "Recovery of parametric models from range images: the case for superquadrics with global deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 2, pp. 131–147, 1990.
- [165] L. Chevalier, F. Jaillet, and A. Baskurt, "Segmentation and superquadric modeling of 3d objects," in *International Conference on Computer Graphics, Visualization and Computer Vision*, 2003.
- [166] E. Bardinet, L. Cohen, and N. Ayache, "Fitting 3-D data using superquadrics and free-form deformations," in *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, vol. 1, pp. 79–83, 1994.
- [167] T. Sederberg and S. Parry, "Free-form deformation of solid geometric model," *SIGGRAPH*, pp. 151–160, 1986.
- [168] Y. Zhang, J. Paik, A. Koschan, and M. Abidi, "3-D object representation from multi-view range data applying deformable superquadrics," in *Proceedings of the 16th International Conference on Pattern Recognition*, vol. 3, pp. 611–614 vol.3, 2002.
- [169] A. Jain and B. Chandrasekaran, *Handbook of Statistics*, ch. Dimensionality and Sample Size Considerations in Pattern Recognition Practice, pp. 835–855. Elsevier Science Pub Co, 1982.
- [170] B. Schoelkopf and A. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2001.

- [171] T. Joachims, *Making large-Scale SVM Learning Practical. Advances in Kernel Methods - Support Vector Learning*. MIT Press, 1999.
- [172] L. Fausett, *Fundamentals of neural networks: architectures, algorithms, and applications*. Prentice-Hall, Inc., 1994.
- [173] C. Debes, C. Weiss, A. Zoubir, and M. Amin, “Wall-clutter mitigation using cross-beamforming in through-the-wall radar imaging,” in *Proceedings of the 18th European Signal Processing Conference (EUSIPCO)*, pp. 1097–1100, 2010.

Curriculum vitae

Name: Ahmed A. Mostafa
Date of birth: 28.01.1976
Place of birth: Alexandria, Egypt
Family status: married

Education

11/1997-09/2004 Arab Academy for science and technology (AAST)
Communication and Electronics Engineering
(Master of Science)
09/1992-09/1997 Arab Academy for science and technology (AAST)
Communication and Electronics Engineering
(Bachelor of Science)
06/2000 High school degree at Mubarak Secondary school
Alexandria, Egypt

Work experience

09/2012 - present Technology Consultant
Soft2tec GmbH, Rüsselsheim, Germany.
11/2008 -09/2012 Research associate at Signal Processing Group
Technische Universität Darmstadt
12/2001 - 10/2008 IT Training manager / Project Manager
AAST Alexandria, Egypt
11/1997 - 03/2000 Instructor, AAST, Alexandria, Egypt

