# A COMPARATIVE ANALYSIS OF DIFFERENT APPROACHES TO TARGET DIFFERENTIATION AND LOCALIZATION USING INFRARED SENSORS

A DISSERTATION SUBMITTED TO

THE DEPARTMENT OF ELECTRICAL AND ELECTRONICS

ENGINEERING

AND THE INSTITUTE OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

By

Tayfun Aytaç

December 2006

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Prof. Dr. Billur Barshan (Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Prof. Dr. Ömer Morgül

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Prof. Dr. Selim Aktürk

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Asst. Prof. Dr. Selim Aksoy

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Asst. Prof. Dr. Ruşen Öktem

Approved for the Institute of Engineering and Science:

Prof. Dr. Mehmet Baray
Director of the Institute

# ABSTRACT

## A COMPARATIVE ANALYSIS OF DIFFERENT APPROACHES TO TARGET DIFFERENTIATION AND LOCALIZATION USING INFRARED SENSORS

Tayfun Aytaç
Ph. D. in Electrical and Electronics Engineering
Supervisor: Prof. Dr. Billur Barshan
December 2006

This study compares the performances of various techniques for the differentiation and localization of commonly encountered features in indoor environments, such as planes, corners, edges, and cylinders, possibly with different surface properties, using simple infrared sensors. The intensity measurements obtained from such sensors are highly dependent on the location, geometry, and surface properties of the reflecting feature in a way that cannot be represented by a simple analytical relationship, therefore complicating the localization and differentiation process. The techniques considered include rule-based, template-based, and neural network-based target differentiation, parametric surface differentiation, and statistical pattern recognition techniques such as parametric density estimation, various linear and quadratic classifiers, mixture of normals, kernel estimator, $k$-nearest neighbor, artificial neural network, and support vector machine classifiers. The geometrical properties of the targets are more distinctive than their surface properties, and surface recognition is the limiting factor in differentiation. Mixture of normals classifier with three components correctly differentiates three types of geometries with different surface properties, resulting in the best performance (100%) in geometry differentiation. For a set of six surfaces, we get a correct differentiation rate of 100% in parametric differentiation based on reflection modeling. The results demonstrate that simple infrared sensors, when coupled with appropriate processing, can be used to extract substantially more information than such devices are commonly employed for. The demonstrated system would find application in intelligent autonomous systems such as mobile robots whose task involves surveying an unknown environment made of different geometry and surface types. Industrial applications where different materials/surfaces must be identified and separated may also benefit from this approach.

*Keywords:* infrared sensors, optical sensing, target differentiation, target localization, surface recognition, position estimation, feature extraction, statistical pattern recognition, artificial neural networks.

# ÖZET

# KIZILBERİSİ ALGILAYICILARLA HEDEF AYIRDETME VE KONUM KESTİRİM YÖNTEMLERİNİN KARŞILAŞTIRMALI İNCELEMESİ

Tayfun Aytaç
Elektrik ve Elektronik Mühendisliği, Doktora
Tez Yöneticisi: Prof. Dr. Billur Barshan
Aralık 2006

Bu çalışma, farklı yüzey özelliklerine sahip düzlem, köşe, kenar ve silindir gibi iç mekanlarda sıkça karşılaşılan öznitelikleri veya hedefleri basit kızılberisi algılayıcılar kullanarak ayırdetme ve konumlandırmaya ilişkin çeşitli tekniklerin başarımlarını karşılaştırmaktadır. Bu tip algılayıcılardan elde edilen yeğinlik ölçümleri hedefin konumuna, geometrisine ve yüzey özelliklerine analitik olarak kolayca ifade edilemeyecek şekilde bağlı olup bu durum ayırdetme ve konumlandırma sürecini zorlaştırmaktadır. Karşılaştırılan teknikler kural-tabanlı, referans sinyallerine dayalı ve yapay sinir ağlarına dayalı hedef ayırdetme, parametrik yüzey ayırdetme ve parametrik yoğunluk kestirimi, farklı doğrusal ve karesel ayırdediciler, karma Gauss sınıflandırıcıları, çekirdek kestiricisi, $k$-en yakın komşuluğu, yapay sinir ağları sınıflandırıcıları ve destekçi vektör makinaları gibi istatistiksel örüntü tanıma tekniklerini içermektedir. Hedeflerin geometrik özellikleri yüzey özelliklerine göre daha ayırdedicidir ve yüzey tipi, ayırdetmede sınırlayıcı etkendir. Üç bileşenli karma Gauss sınıflandırıcıları farklı yüzey özelliklerine sahip üç geometriyi en iyi geometri ayırdetme oranı olarak (%100) doğru ayırdetmektedir. Altı farklı yüzey için yansıma modeline dayalı parametrik ayırdetmede en iyi olarak %100 doğru ayırdetme oranı elde edildi. Sonuçlar, basit kızılberisi algılayıcıların, uygun işlemeyle çok daha fazla bilgi çıkarılarak bilinen yaygın uygulamaları dışında da kullanılabileceğini göstermektedir. Öne sürülen sistem gezgin robotların farklı geometri ve yüzey tiplerinden oluşan bilinmeyen ortamların incelenmesi ve harita çıkarımı gibi uygulamalarda akıllı otonom sistemler tarafından kullanılabilir. Farklı maddelerin/yüzeylerin tanımlanmasının ve ayırdedilmesinin gerektiği endüstriyel uygulamalar da bu yaklaşımdan faydalanabilir.

*Anahtar sözcükler*: kızılberisi algılayıcılar, optik algılama, hedef ayırdetme, hedef

konum kestirimi, yüzey tanıma, konum kestirimi, öznitelik çıkarımı, istatistiksel örüntü tanıma, yapay sinir ağları.

# Acknowledgment

I would like to express my sincere thanks to my thesis supervisor Dr. Billur Barshan for her supervision, guidance, suggestions, and encouragement at all stages during the development of this thesis.

I would also like to thank Dr. Ömer Morgül, Dr. Selim Aktürk, Dr. Selim Aksoy, and Dr. Ruşen Öktem for reading, commenting, and making useful suggestions on my thesis.

It is a pleasure to express my special thanks to Çağrı Yüzbaşıoğlu, for his great friendship, support, patience, and collaboration. I am fortunate enough to be one of his closest friends. Thank you for everything.

Special thanks to Ayşe Yalçın for her friendship, constant support, and encouragement. Thanks for everything she had done for me so far.

I wish to thank all of my friends at the Department of Electrical Engineering, Bilkent University. Special thanks to Uğur Töreyin and Ali Bozbey for sharing their office with me in the last stage of my studies. I want also like to thank the department secretary Mürüvet Parlakay and the laboratory technicians Ersin Başar and Ergün Hırlakoğlu for their help and friendship.

I also want thank my friends, Hale Üstüner and Ebru Bulut, for not leaving me 'alone' in Havelsan.

I am also indebted to my family for their love, patience, and support throughout my life.

# Contents

**1 INTRODUCTION** **1**

1.1 Related Work . . . . . 3

1.2 Contribution . . . . . 5

1.3 Thesis Outline . . . . . 7

**2 INFRARED SENSOR AND THE EXPERIMENTAL SETUP** **8**

2.1 Infrared Sensor . . . . . 8

2.2 Experimental Setup . . . . . 15

**3 RULE-BASED DIFFERENTIATION** **19**

3.1 Differentiation and Localization Algorithm . . . . . 20

3.2 Experimental Verification . . . . . 23

**4 TEMPLATE-BASED DIFFERENTIATION** **26**

4.1 Geometry Differentiation and Localization . . . . . 26

4.1.1 Least-Squares Approach . . . . . 29

4.1.2 Matched Filtering Approach . . . . . 31

4.1.3 Saturated Scans . . . . . 31

4.1.4 Experimental Verification . . . . . 32

4.2 Surface Differentiation and Localization . . . . . 35

4.2.1 Experimental Verification . . . . . 37

4.3 Geometry and Surface Differentiation and Localization . . . . . 41

4.3.1 Experimental Verification . . . . . 43

4.3.2 Limitations of System Performance . . . . . 47

4.3.3 Effect of Orientation of the Targets . . . . . 50

**5 NEURAL NETWORK-BASED DIFFERENTIATION 55**

5.1 ANN Structure and Parameters . . . . . 56

5.2 Differentiation of Geometry Types with ANN . . . . . 60

5.3 Differentiation of Surface Types with ANN . . . . . 62

**6 PARAMETRIC DIFFERENTIATION 66**

6.1 Modeling of Infrared Intensity Scans . . . . . 66

6.2 Experimental Verification . . . . . 72

**7 DIFFERENTIATION BASED ON STATISTICAL PATTERN RECOGNITION TECHNIQUES 81**

7.1 Statistical Pattern Recognition Techniques . . . . . 81

7.1.1 Determination of Geometry . . . . . 86

7.1.2 Determination of Surface Type . . . . . 104

**8 COMPARISON OF THE TECHNIQUES 107**

**9 CONCLUSIONS AND FUTURE WORK 113**

# List of Figures

2.1 (a) Opposed, (b) retroreflective, (c) diffuse, and (d) convergent modes. . . . . 9

2.2 Experimental setup to analyze the effect of various parameters on the performance of the infrared sensor. . . . . 10

2.3 Intensity versus distance characteristics for planar target of different surface properties. . . . . 11

2.4 Effect of surface roughness on the intensity readings for a plane of gray drawing paper. . . . . 12

2.5 Standard deviation versus distance characteristics for various planes. 13

2.6 The mean and the $\pm 25\sigma$ of the intensity measurements versus scan angle for a wooden plane located at $r$ =35 cm and $\theta = 0°$. . . . . 13

2.7 Detectable range of a smooth white plane by the infrared sensors. 14

2.8 Variation of the intensity with respect to distance and angle for a smooth white plane. . . . . 15

2.9 The half-power beamwidth of the infrared sensor. . . . . 16

2.10 (a) The infrared sensor and (b) the experimental setup used in this study. . . . . 17

2.11 Top view of the experimental setup used in target differentiation and localization. The emitter and detector windows are circular with 8 mm diameter and center-to-center separation of 12 mm. (The emitter is above the detector.) Both the scan angle $\alpha$ and the surface azimuth $\theta$ are measured counter-clockwise from the horizontal axis. . . . . 18

2.12 Target primitives used in this study. . . . . 18

3.1 Top view of the experimental setup used in rule-based target differentiation. . . . . 20

3.2 Intensity-versus-scan-angle characteristics for various targets along the line-of-sight of the experimental setup. . . . . 21

4.1 Intensity scans of the four targets at different distances. . . . . 28

4.2 Central intensity versus distance curves for the different geometries. 30

4.3 Intensity scans of the four surfaces at different distances. . . . . 36

4.4 Central intensity versus distance curves for the different surfaces. 37

4.5 Intensity scans for targets (first row, plane; second row, corner; third row, edge) covered with different surfaces (first column, aluminum; second column, white cloth; third column, Styrofoam) at different distances. . . . . 42

4.6 Central intensity (COG) versus distance curves for different targets: (a) plane; (b) corner; (c) edge. . . . . 44

4.7 Intensity scans for a wooden (a) corner at 65 cm, (b) edge at 35 cm for orientations between 0° and 35° with 2.5° increments. The curves with the dotted lines indicate 0° orientation. . . . . 53

5.1 Activation function used in the neural networks. . . . . 56

5.2 Neural network structure used in the study. . . . . 57

5.3 Test and training errors while pruning the network with Optimal Brain Surgeon. . . . . 62

5.4 Neural network after pruned with Optimal Brain Surgeon. . . . . 63

6.1 Lambertian (diffuse) reflection from an opaque surface. Note how the intensity decreases with increasing $\alpha$ but is of equal magnitude in every direction. . . . . 67

6.2 Specular reflection from an opaque surface. . . . . 68

6.3 Intensity scans of the eight surfaces collected between 30 to 52.5 cm in 2.5 cm increments. Solid lines indicate the model fit and the dotted lines indicate the experimental data. . . . . 71

6.4 Variation of the parameters (a) $C_0$, (b) $C_1$, and (c) $z$ with respect to the maximum intensity of the scan. . . . . 73

7.1 Intensity scans of the planes covered with seven planar surfaces collected at different ranges [see Figure 7.4(c)]. Solid lines indicate the model fit and the dotted lines indicate the actual data. . . . . 83

7.2 Intensity scans of the edges covered seven surfaces collected at different ranges [see Figure 7.4(c)]. Solid lines indicate the model fit and the dotted lines indicate the actual data. . . . . 84

7.3 Intensity scans of the cylinders covered with seven surfaces collected at different ranges [see Figure 7.4(c)]. Solid lines indicate the model fit and the dotted lines indicate the actual data. . . . . 85

7.4 Variation of the parameters (a) $C_0$, (b) $C_1$, and (c) $z$ with respect to maximum intensity (dashed, dotted, and solid lines are for planes, edges, and cylinders, respectively). . . . . 87

7.5 Discriminant functions for PDE when the $[C_0, I_{\max}]$ feature vector is used. . . . . 88

7.6 2-D normal contour plots for (a) homoscedastic (b) heteroscedastic PDE when the $[C_1, I_{\max}]^T$ feature vector is used. . . . . 89

7.7 Discriminant functions for PDE when the $[C_1, I_{\max}]^T$ feature vector is used. . . . . 90

7.8 Discriminant functions for the MoN classifier when the $[C_1, I_{\max}]^T$ feature vector is used. . . . . 93

7.9 Discriminant functions for the KE and the $k$-NN classifier when the $[C_1, I_{\max}]^T$ feature vector is used. . . . . 99

7.10 Discriminant functions for ANN classifiers when the $[C_1, I_{\max}]^T$ feature vector is used. . . . . 101

# List of Tables


3.1 Confusion matrix (P: plane, C: corner, E: edge, CY: cylinder). . . 23

3.2 Performance parameters of the algorithm. . . . . 24

4.1 Confusion matrix: least-squares based classification (max/dip variation). . . . . 33

4.2 Confusion matrix: least-squares based classification (COG variation). . . . . 33

4.3 Confusion matrix: matched filter based classification. . . . . 34

4.4 Absolute range and azimuth estimation errors over all test targets. (LS: least-squares, MF: matched filter.) . . . . . 34

4.5 Confusion matrix: least-squares based recognition (maximum intensity variation). (AL: aluminum, WW: white wall, BR: brown paper, ST: Styrofoam). . . . . 38

4.6 Confusion matrix: least-squares based recognition (COG variation). 38

4.7 Confusion matrix: matched filtering based recognition. . . . . 39

4.8 Absolute range and azimuth estimation errors over all surfaces. . . 40

4.9 Confusion matrix: least-squares based classification (maximum variation) (WC: white cloth). . . . . 43

4.10 Confusion matrix: least-squares based classification (COG variation). . . . . 45

4.11 Confusion matrix: matched filter based classification. . . . . 46

4.12 Absolute range and azimuth estimation errors over all test targets. 47

4.13 Confusion matrix for planar targets with unfamiliar surface. (WO: wood, VI: violet paper, BL: black paper, WH: white paper.) . . . 48

4.14 Confusion matrix for corner targets with unfamiliar surface. . . . 49

4.15 Confusion matrix for edge targets with unfamiliar surface. . . . . 50

4.16 Confusion matrix for cylindrical targets with familiar surface. . . 51

4.17 Confusion matrix for cylindrical targets with unfamiliar surface. . 52

5.1 Confusion matrix for ANN before Optimal Brain Surgeon: results are outside (inside) the parentheses for maximum intensity (COG) based azimuth estimation. . . . . 61

5.2 Confusion matrix for ANN after Optimal Brain Surgeon. . . . . 63

5.3 Confusion matrix for three geometries and three surface types. . . 64

6.1 Confusion matrix: $C_1$-based differentiation (initial range to the surface is estimated using the maximum intensity of the scan). . . 75

6.2 Confusion matrix: $C_1$-based differentiation (initial range to the surface is estimated using the maximum intensity of the scan). . . 76

6.3 Confusion matrix: $C_1$-based differentiation (initial range to the surface is estimated using the maximum intensity of the scan). . . 77

6.4 Confusion matrix: $C_1$-based differentiation (range to the surface is known). . . . . 78

6.5 Confusion matrix: $C_1$-based differentiation (range to the surface is known). . . . . 79

6.6 Confusion matrix: $C_1$-based differentiation (range to the surface is known). . . . . 80

6.7 Confusion matrix: $C_1$-based differentiation (initial range estimate is taken as half of the operating range for all surfaces). . . . . 80

7.1 Confusion matrix: homoscedastic PDE using the $[C_0, I_{\max}]^T$ feature vector. Numbers outside (inside) the parentheses are for the training (test) scans. . . . . 86

7.2 Confusion matrix: heteroscedastic PDE using the $[C_0, I_{\max}]^T$ feature vector. . . . . 88

7.3 Confusion matrix: homoscedastic PDE using the $[C_1, I_{\max}]^T$ feature vector. . . . . 91

7.4 Confusion matrix: heteroscedastic PDE using the $[C_1, I_{\max}]^T$ feature vector. . . . . 91

7.5 Confusion matrix: MoN classifier ($M = 3$) using the $[C_1, I_{\max}]^T$ feature vector. . . . . 94

7.6 Confusion matrix: linear classifier by KL expansion of the common covariance matrix. . . . . 94

7.7 Confusion matrix: logistic linear classifier. . . . . 95

7.8 Confusion matrix: Fisher's least-squares linear classifier. . . . . 96

7.9 Confusion matrix: quadratic discriminant classifier. . . . . 97

7.10 Confusion matrix: ANN trained with BP. . . . . 100

7.11 Confusion matrix: ANN trained with LM. . . . . 101

7.12 Confusion matrix: ANN trained with LP. . . . . 102

7.13 Correct differentiation percentages for different classifiers (PDE-HM: Parametric density estimation-homoscedastic, PDE-HT: Parametric density estimation-heteroscedastic, LC-KL: Linear classifier-Karhunen L*ó*eve, LC-LOG: Linear classifier-logistic, LC-FIS: Linear classifier-Fisher's least-squares, NM: nearest mean classifier, NMS: nearest mean scaled classifier, QC: quadratic classifier, MoN-2: Mixture of normals with two components, MoN-3: Mixture of normals with three components, KE: kernel estimator, $k$-NN: $k$-nearest neighbor, ANN-BP: ANN trained with BP, ANN-LM: ANN trained with LM, ANN-LP: ANN trained with LP, SVM-P: SVM with polynomial kernel, SVM-E: SVM with exponential kernel, SVM-R: SVM with radial kernel). . . . . 105

8.1 Overview of the differentiation techniques compared (U: used, S: stored, and NS: not stored). . . . . 110

# Chapter 1

# INTRODUCTION

Target differentiation is of considerable interest for intelligent systems that need to interact with and operate in their environment autonomously. Such systems rely on sensor modules which are often their only available source of information. Since the resources of such systems are limited, the available resources should be used in the best way possible. It is desirable to maximally exploit the capabilities of lower cost sensors before more costly and sophisticated sensors with higher resolution and higher resource requirements are employed. This can be achieved by employing better characterization and physical modeling of these sensors.

Although ultrasonic sensors have been widely used for object detection and ranging [1–6], they are limited by their large beam-width and the difficulty of interpreting their readings due to specular, higher-order, and multiple reflections from surfaces. Furthermore, many readily available ultrasonic systems cannot detect objects up to 0.5 m which corresponds to their blank- out zone. Therefore, in performing tasks at short distances from objects, use of inexpensive and widely available sensors such as simple infrared detectors are preferable to employing ultrasonic sensors or more costly laser and vision systems. Furthermore, in a sensor-fusion framework, infrared sensors would be perfectly complementary to these systems which are not suitable for close-range detection. Infrared detectors offer faster response times and better angular resolution than ultrasonic sensors

and provide intensity readings at nearby ranges (typically from a few centimeters up to a meter). The intensity versus range characteristics are nonlinear and dependent on the properties of the surface and environmental conditions. Consequently, a major problem with the use of simple infrared detectors is that it is often not possible to make accurate and reliable range estimates based on the value of a single intensity return because the return depends on both the geometry and surface properties of the encountered object. Likewise, the surface properties and the geometry of the target cannot be deduced from simple intensity returns without knowing its position and orientation.

Due to single intensity readings not providing much information about the target properties, recognition capabilities of infrared sensors have been underestimated and underused in most work. To achieve accurate results with these sensors, their nonlinear characteristics should be analyzed and modeled based on experimental data. Armed with such characterization and modeling, their potential can be more fully exploited and their usage can be extended beyond simple tasks such as counting and proximity detection. The aim of this study is to maximally realize the potential of these simple sensors so that they can be used in more complicated tasks such as differentiation, recognition, clustering, docking, perception of the environment and surroundings, and map building. With the approaches considered in this thesis, we can differentiate a moderate number of targets and/or surfaces commonly encountered in indoor environments, using a simple infrared system consisting of one emitter and one detector. We provide a comparison of these approaches based on real data acquired from simple infrared sensors. The results indicate that if the data acquired from such simple infrared sensors are processed effectively through the use of suitable techniques, substantially more information about the environment can be extracted than is commonly achieved with conventional usage.

## 1.1 Related Work

The use of infrared sensors in the pattern recognition area has been mostly limited to the recognition or detection of features or targets in conventional two-dimensional images. Examples of work in this category include face identification [7], automatic target recognition [8], target tracking [9], automatic vehicle detection [10], remote sensing [11], detection and identification of targets in background clutter [12, 13], and automated terrain analysis [14]. We note that the position-invariant target differentiation and position estimation achieved in this thesis are different from such operations performed on conventional images [15, 16] in that here we work not on direct "photographic" images of the targets obtained by some kind of imaging system, but rather on angular intensity scans obtained by rotating a point sensor. The targets we differentiate are not patterns in a two-dimensional image whose coordinates we try to determine, but rather objects in space, exhibiting depth, whose position with respect to the sensing system we need to estimate. For this reason, position-invariant pattern recognition and position estimation achieved in this thesis is different from such operations performed on conventional images [15–25].

Application areas of infrared sensing include robotics and automation, process control, remote sensing, and safety and security systems. More specifically, they have been used in simple object and proximity detection [26], counting [27], distance and depth monitoring [28], floor sensing, position control [29], obstacle/collision avoidance [30], and machine vision systems [31]. Infrared sensors are used in door detection [32], mapping of openings in walls [33], as well as monitoring doors/windows of buildings and vehicles, and light curtains for protecting an area. In [34], an automated guided vehicle detects unknown obstacles by means of an "electronic stick" consisting of infrared sensors, using a strategy similar to that adopted by a blind person. In [35], infrared sensors are employed to locate edges of doorways in a complementary manner with sonar sensors. Other researchers have also dealt with the fusion of information from infrared and sonar sensors [36, 37, 38] and infrared and radar systems [39, 40]. In [26], infrared proximity sensing for a robot arm is discussed. Following this work, [30] describes a

robot arm completely covered with an infrared skin sensor to detect nearby objects. Processing the data from the artificial infrared skin by motion planning algorithms, real-time collision avoidance for the entire arm body is achieved in an unknown or dynamic environment.

In [41], the properties of a planar surface at a known distance have been determined using the Phong illumination model, and using this information, the infrared sensor employed has been modeled as an accurate range finder for surfaces at short ranges. Reference [42] also deals with determining the range of a planar surface. By incorporating the optimal amount of additive noise in the infrared range measurement system, the authors were able to improve the system sensitivity and extend the operating range of the system. A number of commercially available infrared sensors are evaluated in [43]. References [44, 45] describe a passive infrared sensing system which identifies the locations of the people in a room. Infrared sensors have also been used for automated sorting of waste objects made of different materials [46, 47]. In [48], an infrared sensor-based system which can measure distances up to 1 m is described. References [49, 50, 51] deal with optical determination of depth information. In [52], simulation and evaluation of the recognition abilities of active infrared sensor arrays are considered for autonomous systems using a ray-tracing approach.

In earlier work [53], the authors developed a novel range estimation technique using an infrared emitter-detector pair which is independent of surface type since it is based on the position of the maximum intensity value instead of surface-dependent absolute intensity values. An intelligent feature of the system is that its operating range is made adaptive based on the maximum intensity of the detected signal.

In the thesis work described by [54], infrared sensors are used for position estimation. Reflectance from spherical objects is modeled by considering the position, orientation, and the characteristics of the emitter and detector, the position, size, and reflectivity of the spherical object, and the intensity of the reflected light. 3-D position estimation of objects have been implemented using a

non-touch screen. 2-D object position estimation is implemented using an electrically powered wheelchair, whose movement is controlled by the head movement through infrared sensors.

## 1.2 Contribution

In this thesis, we propose several new techniques for processing infrared intensity signals and compare their performances with several existing approaches for differentiation and localization of commonly encountered features in indoor environments. The classification approaches include rule-based, template-based (*matched filter* and *least-squares* variations), neural network-based differentiation, parametric differentiation, and pattern recognition techniques such as maximum-likelihood estimation, various linear and quadratic classifiers, mixture of normals, $k$-nearest neighbor classifier, neural network classifier, and support vector machine classifier. One advantage of our system is that it does not greatly depend on environmental conditions, since we employ an *active* sensing modality.

To the best of our knowledge, no attempt has been made to differentiate and estimate the position of several kinds of targets using infrared sensors. Also, a comparative study based on experimental work does not exist for target differentiation using infrared sensors. One of the major contributions of this thesis is that it provides such a comparison. The results indicate that it is possible to extract a significantly greater amount of information from simple optical sensors than is commonly achieved with conventional usage.

As a first attempt to differentiation of targets with simple infrared sensors, we employed a rule-based approach which is based on extracting empirical rules by inspecting the nature of the infrared intensity scans. For this purpose, angular intensity scans are obtained from two infrared sensors horizontally mounted on a rotary table. The method can achieve position-invariant target differentiation without relying on the absolute return signal intensities of the infrared sensors. The target primitives employed in the rule-based approach are plane, corner, edge,

and cylinder, all made of unpolished oak wood. Based on tests with experimental data, an average correct differentiation rate of 91.3% is achieved.

The template-based approach is based on comparing the acquired infrared intensity scans obtained from targets located at randomly selected distance and azimuth values with previously stored templates acquired from targets located at predetermined distances and the line-of-sight of the experimental setup. Hence, this approach relies on the distinctive natures of the infrared intensity scans and requires the storage of a complete set of reference scans of interest. We considered the following three different cases: targets with different geometrical properties but made of the same surface material (97% correct differentiation rate), targets made of different surface materials but of the same planar geometry (87% correct differentiation rate), and targets with both different geometry and surface properties (80% correct differentiation rate).

As an alternative to template-based differentiation, artificial neural networks are proposed for geometry and surface type determination. The training algorithms employed are back-propagation (BP) and Levenberg-Marquardt (LM). The networks trained with LM are pruned with Optimal Brain Surgeon technique [55] for the optimal network structure. Pruning also results in improved classification. The differentiation results are comparable with those obtained with template-based target differentiation, where geometry type of the targets is classified with 99% accuracy and an overall correct differentiation rate of 78.4% is achieved for all surfaces.

The parametric approach is based on modeling of infrared intensity scans. In parametric surface differentiation, only the reflection coefficients obtained using the proposed reflection model are considered as parameters and used in the differentiation process, instead of using the complete infrared intensity scans as in the previous differentiation approaches. The surface materials considered are unpolished oak wood, Styrofoam packaging material, white painted matte wall, white and black cloth, and white, brown, and violet paper (matte). For a set of six surfaces including Styrofoam packaging material, white painted matte wall, white or black cloth, and white, brown, and violet paper (also matte), we got a

correct differentiation rate of 100%.

We extended the parametric surface differentiation approach to differentiation of both the geometry and surface type of the targets using statistical pattern recognition techniques. We constructed feature vectors based on the parameters obtained modeling of angular infrared intensity scans from different targets to determine their geometry and/or surface type. The techniques considered in this thesis include statistical pattern recognition techniques (parametric density estimation, Karhunen Lóeve based classifier, logistic linear classifier, Fisher's linear classifier, nearest mean classifier and its scaled version, quadratic discriminant classifier, mixture of normals, kernel estimator, $k$-nearest neighbor, artificial neural network, and support vector machine classifiers). Mixture of normals classifier with three components correctly differentiates three types of geometries with different surface properties, resulting in the best performance (100%) in geometry differentiation.

## 1.3 Thesis Outline

This thesis is organized as follows: Chapter 2 describes the infrared sensor and the experimental setup used in this study. In Chapter 3, rule-based target differentiation [56] is explained. Chapter 4 describes template-based geometry [57], surface [58], and both geometry and surface differentiation and localization and discusses the limits of the proposed approaches [59]. As an alternative to template-based differentiation, artificial neural network-based geometry and surface type determination is proposed in Chapter 5. Chapter 6 provides differentiation of planar surfaces based on parametric modeling of the infrared intensity scans [60]. In Chapter 7, statistical pattern recognition techniques using reflection coefficients are proposed for geometry and surface type determination [61]. A comparison of the techniques is provided in Chapter 8. Concluding remarks are made and directions for future work are provided in Chapter 9.

# Chapter 2

# INFRARED SENSOR AND THE EXPERIMENTAL SETUP

We believe that for proper operation of a sensor, the parameters affecting its operation should be thoroughly investigated. In this chapter, the effects of parameters such as range, azimuth, and surface properties on the operation of the infrared sensor are investigated.

This chapter is organized as follows: The operation of the infrared sensor and the parameters affecting its operation are investigated thoroughly in Section 2.1. The experimental setup is described in Section 2.2.

## 2.1 Infrared Sensor

The emitter-detector configuration of infrared sensors can be classified into four groups as opposed, retroreflective, diffuse, and convergent modes [31] (Figure 2.1). Opposed mode is used, for instance, in remote controls. The retroreflective mode, in which the emitted energy is reflected from a retroreflector, such as a corner cube is commonly used in, for instance, doorway detectors in buildings and elevator doors. It is also used for reference marking purposes in automated guided vehicles.

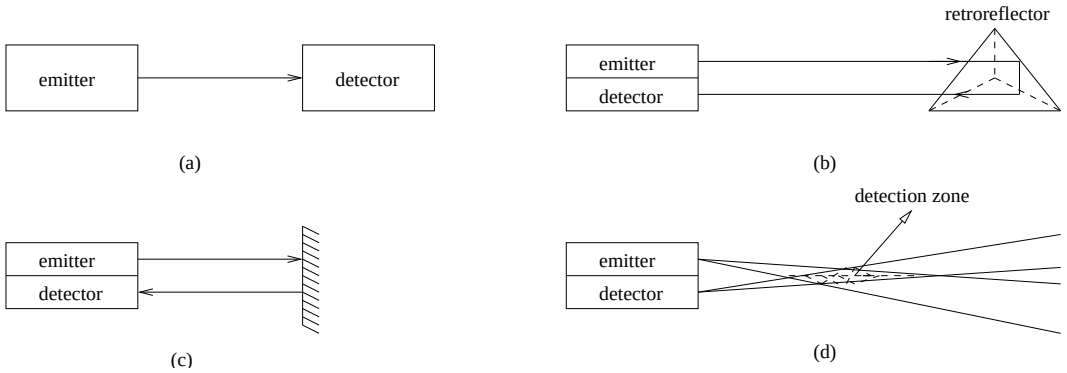



Figure 2.1: (a) Opposed, (b) retroreflective, (c) diffuse, and (d) convergent modes.

Mostly used in object detection is the diffuse mode, where the emitted energy is reflected from the object of interest. In the convergent mode, the optical axis of the emitter-detector is tilted in order to detect objects over a specific range.

In our experimental work, the IRS-U-4A infrared sensor [62] is used (see Figure 2.2). The sensor works with 20–28 V DC input voltage, and provides an analog output voltage proportional to the measured intensity reflected off the target. The detector window is covered with an infrared filter to minimize the effect of ambient light on the intensity measurements. Indeed, when the emitter is turned off, the detector reading is essentially zero. The sensitivity of the device can be adjusted with a potentiometer to set the range of operation of the system.

Various surfaces with different colors and surface properties have been considered. To analyze the effect of the surface roughness, packing materials with different reflection properties are employed. The experimental setup used for this purpose is shown in Figure 2.2, where a planar surface is employed for the purpose of uniform characterization of different surfaces. The plane is chosen large enough to contain the infrared spot size. The optical axis of the infrared sensor is coincident with the normal of the plane. Measurements are taken with the potentiometer adjusted both at its rightmost and leftmost positions, corresponding to minimum (5 cm) and maximum range of operation (70 cm), respectively.

To study the effect of target range, azimuth, and surface parameters on the measurements, intensity samples are acquired for each position and surface, and

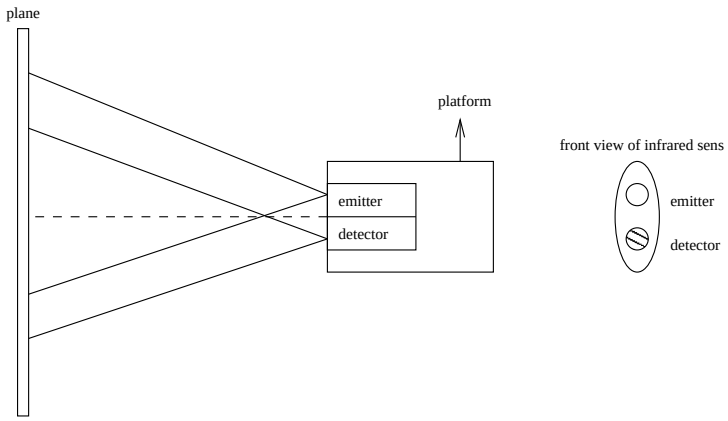


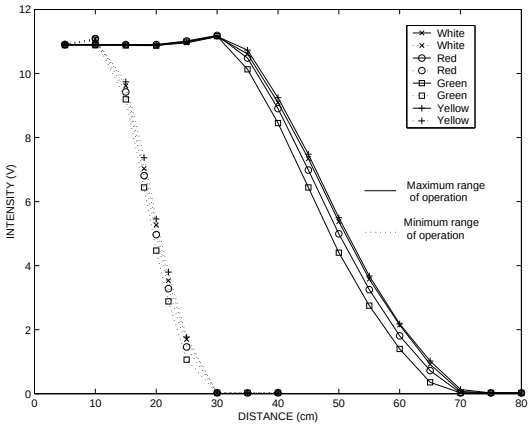

Figure 2.2: Experimental setup to analyze the effect of various parameters on the performance of the infrared sensor.

their mean and standard deviations are calculated. In Figure 2.3(a), the plots of intensity versus distance are given for the plane covered with white, red, green, and yellow copier/printer papers. Notice that for each color, there is a certain range of operation determined by saturation at the lower end and loss of signal at the higher end (beyond a certain range, the output voltage is not detectable). For the situation where the potentiometer is adjusted at its rightmost position, it is possible to deduce the range of the plane of different colors within a few centimeters error. We observe that the color does not have a strong effect on the output intensity which makes the system suitable for range detection of different colored surfaces.

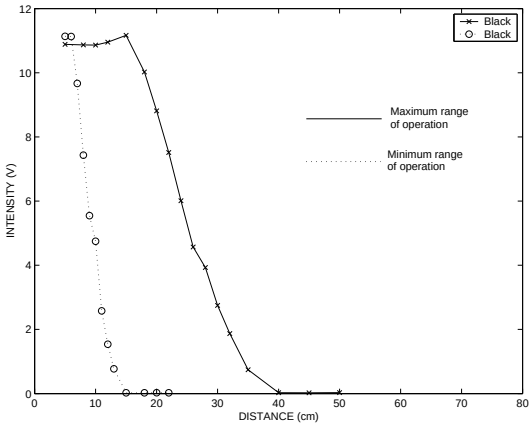
Unlike the planes above, the plane covered with glossy, smooth, black plane (craft paper) showed different behavior due to its high absorption property (Figure 2.3(b)).

Drawing papers having gray, dark blue, and brown colors are also employed. These papers are slightly thicker than copier papers and have a little more roughness on one side than the other. Because of their different surface properties, their characteristics differ from those of the copier papers. The intensity variations with respect to distance are given in Figure 2.3(c).

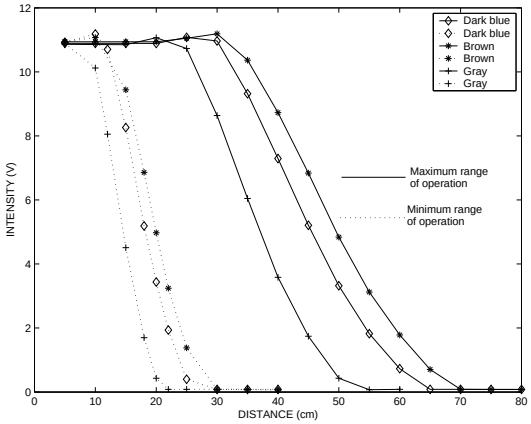
Blister packaging materials made of transparent colorless nylon with large and



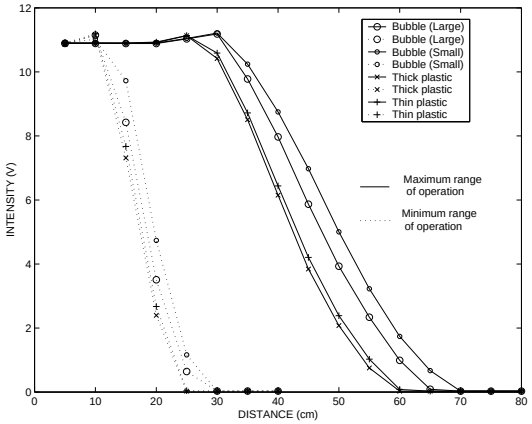

(a) white, red, green, and yellow copier/printer papers




(b) black craft paper




(c) drawing papers of different colors




(d) various packing materials

Figure 2.3: Intensity versus distance characteristics for planar target of different surface properties.

small bubbles and Styrofoam packaging materials are also used to investigate the effect of different surface properties on the measurements. The blister packaging material with small bubbles has a honeycomb pattern of uniformly distributed circular bubbles of diameter 1.0 cm and height 0.3 cm, with a center-to-center separation of 1.2 cm. The blister packaging material with large bubbles has the same pattern with diameter, height, and center-to-center separation of 2.5, 1.0, and 2.8 cm, respectively. The variation of the intensity with respect to distance is given in Figure 2.3(d). The Styrofoam packaging material absorbs more energy than the blister packaging materials. As expected, for a given distance, the return signal for the plane with small bubbles is greater than that with large bubbles.

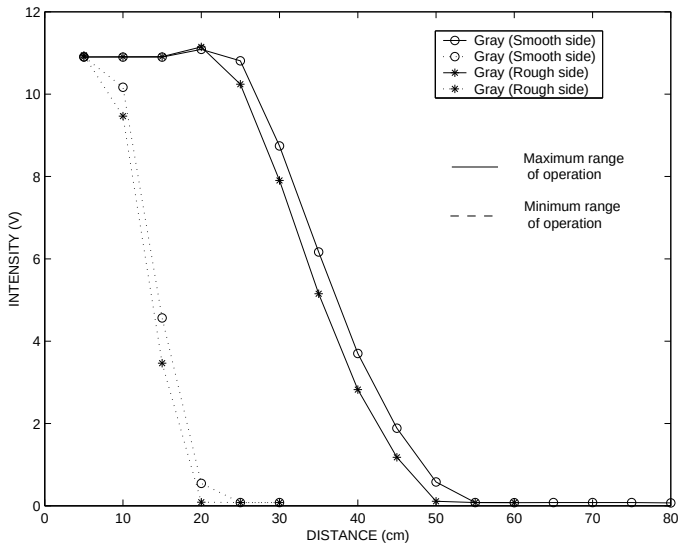



Figure 2.4: Effect of surface roughness on the intensity readings for a plane of gray drawing paper.

This is the result of enhanced multi-directional reflection due to large bubbles.

In Figure 2.4, the results obtained with both sides of the gray drawing paper are displayed, one surface being slightly rougher than the other. As seen from the graph, the surface roughness may result in erroneous readings even for a plane of the same color.

The variation of the standard deviation with respect to distance for various planes is given in Figure 2.5. For a given distance value and a surface type, the standard deviation was calculated over 10,000 intensity measurements. The standard deviation varies approximately within a band of 0.04 $\pm$0.01 V.

The variation of the standard deviation with respect to the scan angle is illustrated in Figure 2.6 for a wooden plane located at $r$ =35 cm and $\theta = 0°$. The mean and the standard deviation values of the scan were calculated over 1,000 intensity measurements at each step of the scan. Figure 2.6 illustrates the mean value $\pm 25\sigma$. The standard deviation was calculated to vary between a minimum value of 0.006 V and a maximum value of 0.04 V.

Now, we turn our attention to the problem of determining the operating range

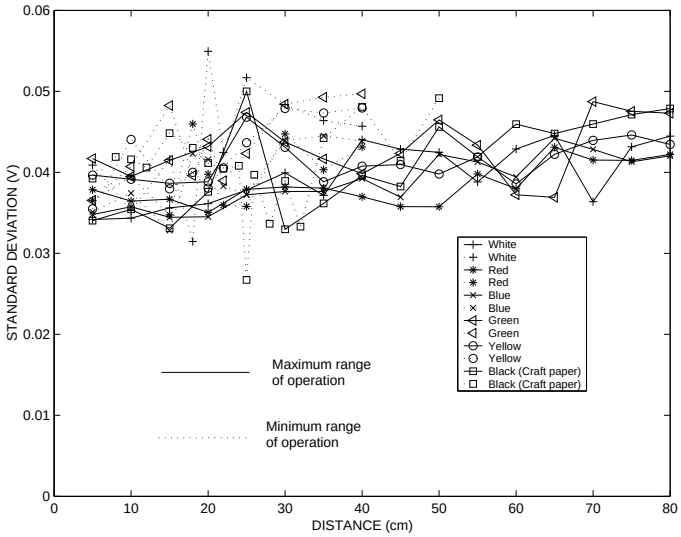



Figure 2.5: Standard deviation versus distance characteristics for various planes.

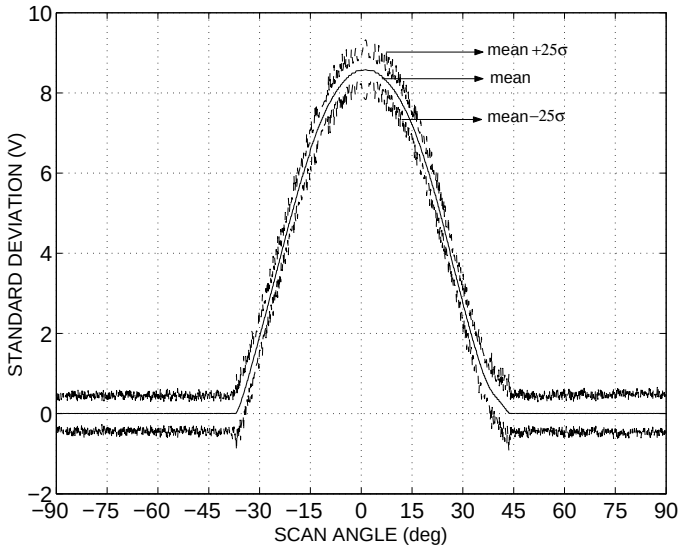



Figure 2.6: The mean and the $\pm 25\sigma$ of the intensity measurements versus scan angle for a wooden plane located at $r$ =35 cm and $\theta = 0°$.

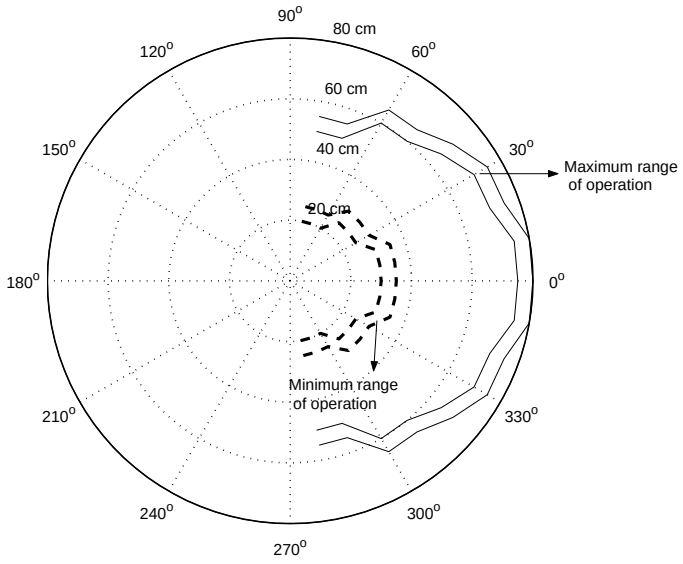



Figure 2.7: Detectable range of a smooth white plane by the infrared sensors.

and angle of our system. To this end, the sensing unit will be situated on the grid points shown in Figure 2.7, in each case pointing towards the center of the radial grid. We have considered both of the extreme settings of the potentiometer. Using the plane covered with white copier/printer paper, measurements are taken at 5 cm intervals from 5 to 80 cm, and at $\theta = 10°$ intervals from $\theta = 0°$ and $\theta = 80°$ with the normal of the plane (smooth, white plane is chosen to minimize the effect of the diffuse reflectance ratios [63]).

The variation of the intensity with respect to distance and angle for the white plane is given in Figure 2.8. By using these plots, the detectable range of the plane is given in Figure 2.7. The outer curve is composed of points whose intensities are less than 0.1 V, and the inner curve is composed of points whose intensities are greater than or equal to 0.1 V. The curves are given both for the rightmost (solid lines) and leftmost (dashed lines) positions of the potentiometer. For the rightmost position of the potentiometer, the infrared sensor can detect the plane making $\theta = 80°$ angle with the normal of the plane at 50 cm. On the other hand, at the same angle, the infrared sensor can detect the plane at 20 cm at the leftmost position of the potentiometer. As seen from the plot, the intensity

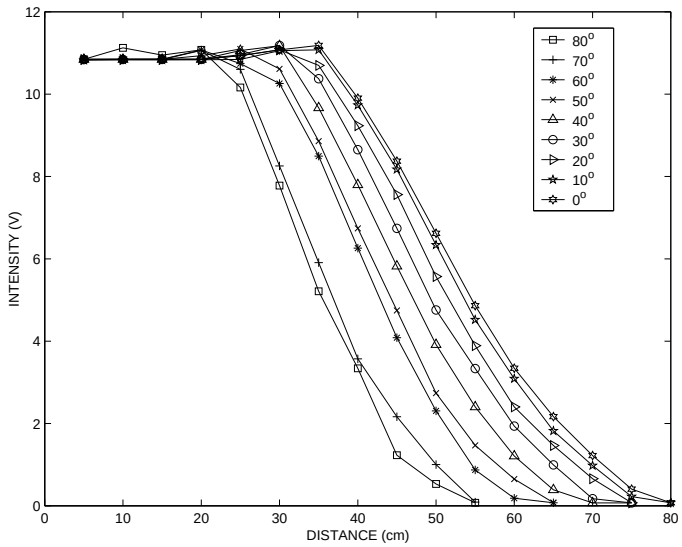



Figure 2.8: Variation of the intensity with respect to distance and angle for a smooth white plane.

depends on the position of the plane with respect to the infrared sensor. As the line-of-sight of the infrared sensor deviates from the normal of the plane, the intensity decreases (Figure 2.8).

The half-power beamwidth of the infrared sensor is found as in [64] by setting the intensity to $1/\sqrt{2}$ of the maximum reading obtained. The half-power beamwidth is found to be approximately $\theta = 3.3^{\circ}$ (Figure 2.9), which makes it useful for object detection due to its acceptable angular resolution.

## 2.2 Experimental Setup

The infrared sensor is mounted on a 12 inch rotary table [65] to obtain angular intensity scans from these targets. The close-up view of the infrared sensor and the photograph of the experimental setup can be seen in Figure 2.10. The schematics of the experimental setup is shown in Figure 2.11. For the rule-based classification, described in the next chapter, we use two infrared sensors horizontally mounted on the rotary table with a center-to-center separation of 11 cm

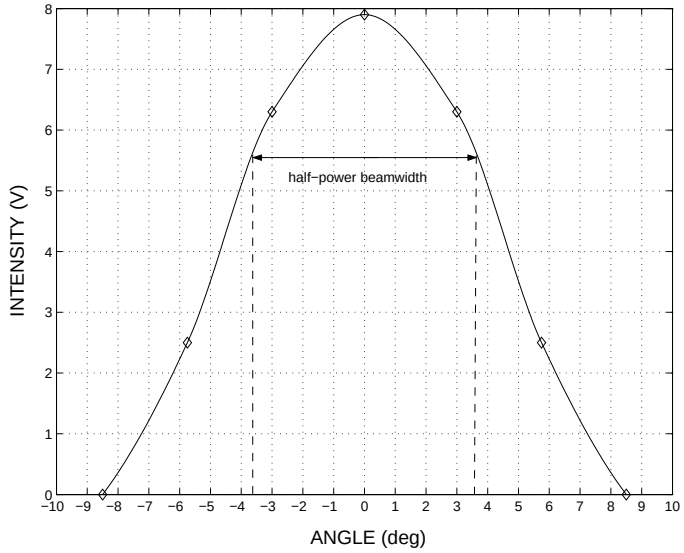
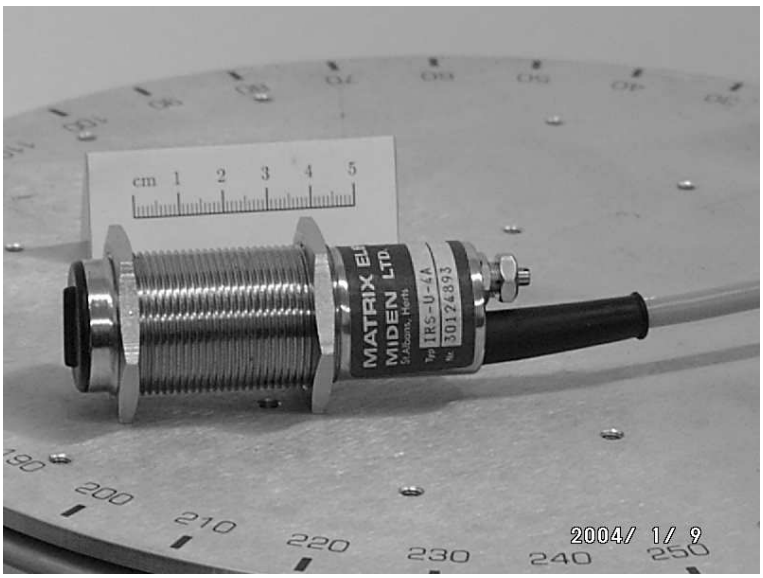


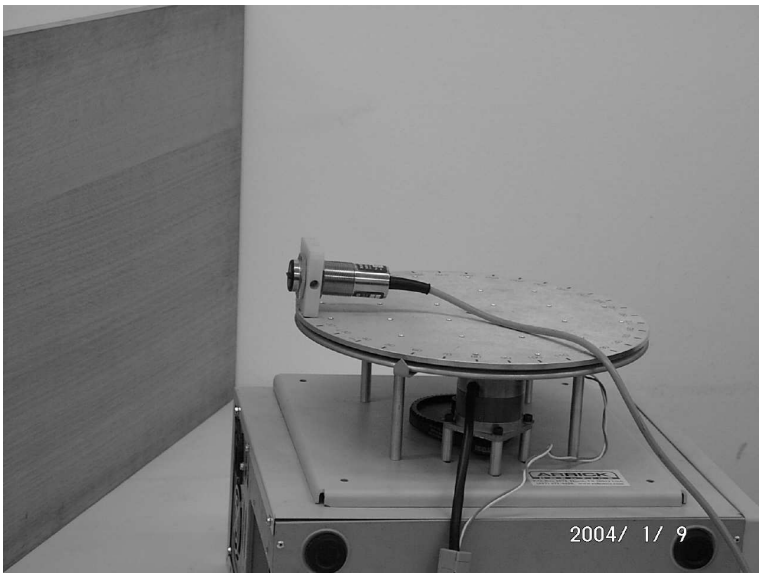

Figure 2.9: The half-power beamwidth of the infrared sensor.

(see Figure 3.1). The target primitives employed in this study are a plane, a 90° corner, a 90° edge, and a cylinder of radius 4.8 cm, whose cross-sections are given in Figure 2.12. The horizontal extent of all targets other than the cylinder is large enough that they can be considered infinite and thus edge effects need not be considered. They are covered with different materials of different surface properties, each with a height of 120 cm. For the methods discussed in this study, results will be given for targets of different geometry and/or surface properties and their combinations.

In this chapter, we discussed the effects of range, azimuth, and surface properties on the operation of the infrared sensors and introduced the experimental setup. In the following chapters, we will describe and compare different methods for target differentiation and localization.




(a)




(b)

Figure 2.10: (a) The infrared sensor and (b) the experimental setup used in this study.

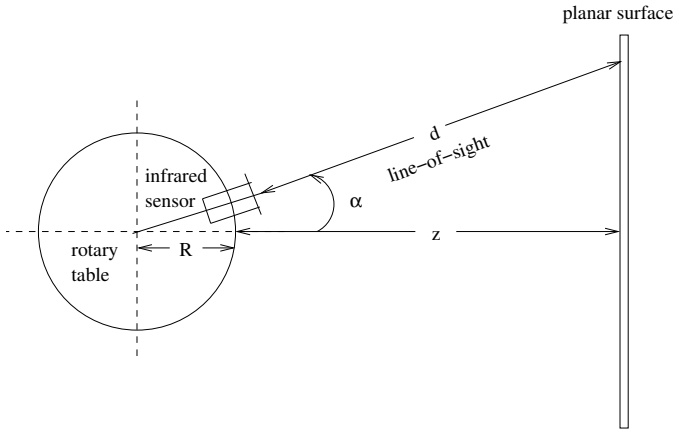



Figure 2.11: Top view of the experimental setup used in target differentiation and localization. The emitter and detector windows are circular with 8 mm diameter and center-to-center separation of 12 mm. (The emitter is above the detector.) Both the scan angle $\alpha$ and the surface azimuth $\theta$ are measured counter-clockwise from the horizontal axis.

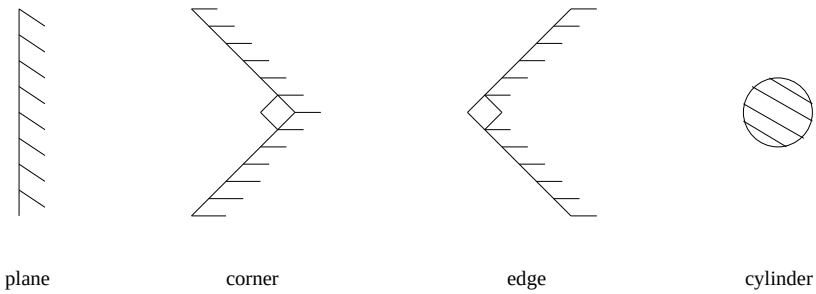



Figure 2.12: Target primitives used in this study.

# Chapter 3

# RULE-BASED DIFFERENTIATION

In this chapter, we consider processing information from a pair of infrared sensors using a rule-based approach for target differentiation and localization. The work in this chapter was published in [56]. The advantages of a rule-based approach are shorter processing times, greater robustness to noise, and minimal storage requirements in that it does not require storage of any reference scans: the information necessary to differentiate the targets is completely embodied in the decision rules [66]. Examples of related approaches with ultrasonic sensors may be found in [67, 68].

Our method is based on angularly scanning of the target over a certain angular range. We use two infrared sensors horizontally mounted on a 12 inch rotary table [65] with a center-to-center separation of 11 cm [Figure 3.1] to obtain angular scans $I(\alpha)$ from the targets. Targets are scanned from $-60^\circ$ to $60^\circ$ in $0.15^\circ$ increments, and the mean of 100 samples are calculated at each position of the rotary table. The targets are situated at ranges varying between 20 and 65 cm. The outputs of the infrared sensors are multiplexed to the input of an 8-bit microprocessor compatible analog-to-digital converter chip having a conversion time of 100 $\mu$sec.

## 3.1 Differentiation and Localization Algorithm

Some sample scan patterns obtained from the targets are shown in Figure 3.2. Based on these patterns, it is observed that the return signal intensity patterns for a corner, which have two maxima and a single minimum (a double-humped pattern), differ significantly from those of other targets which have a single maximum [Figure 3.2(b)]. The double-humped pattern is a result of the two orthogonal planes constituting the corner. Because of these distinctive characteristics, the corner differentiation rule is employed first. We check if the scan pattern has two humps or not. If so, it is a corner. The average of the angular locations of the dips in the middle of the two humps for the left and right infrared sensors provides an estimate of the angular location of the corner.

If the target is found not to be a corner, we next check whether it is a plane or not. As seen in Figure 3.2(a), the difference between the angular locations of the maximum readings for the planar targets is significantly smaller than for other targets. Planar targets are differentiated from other targets by examining the absolute difference of the angle values at which the two intensity patterns have their maxima. If the difference is less than an empirically determined reference value, then the target is a plane; otherwise, it is either an edge or a cylinder. (In the experiments, we have used a reference value of 6.75°.) The azimuth estimation of planar targets is accomplished by averaging the angular locations of the maxima of the two scans associated with the two sensors.

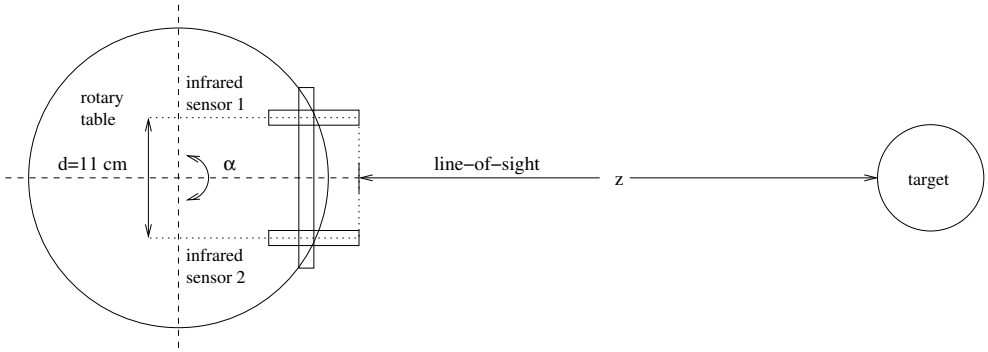
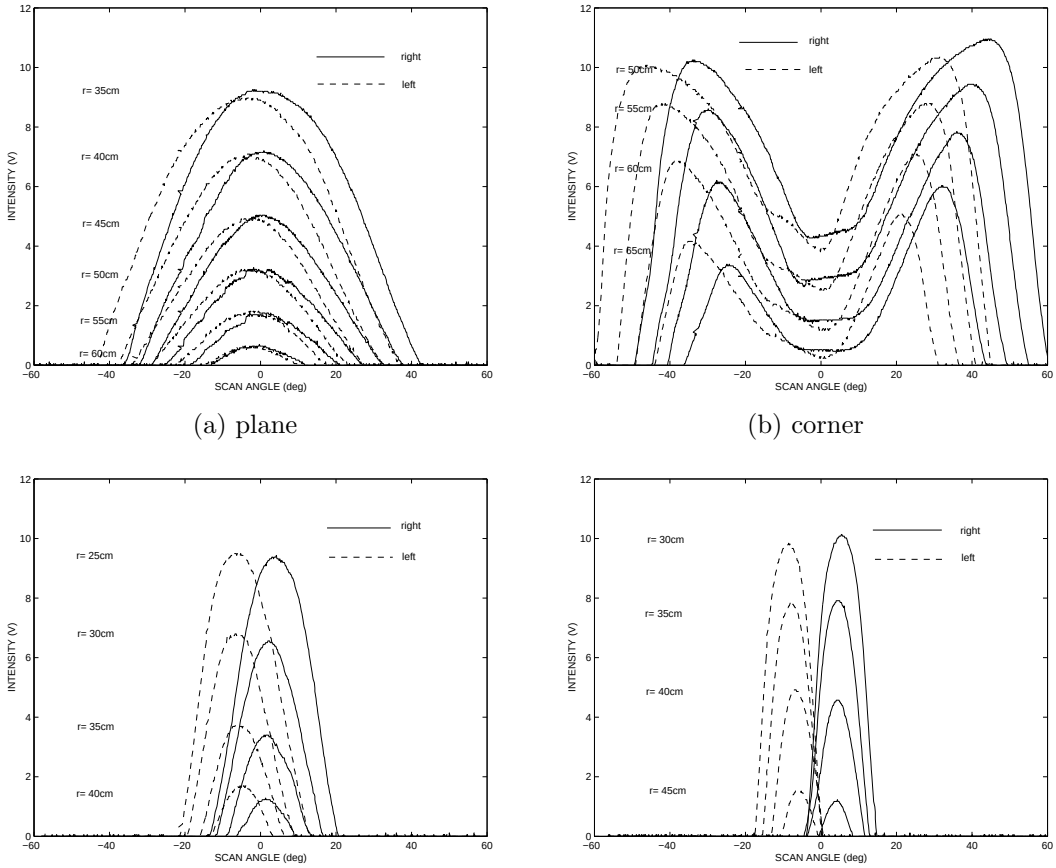



Figure 3.1: Top view of the experimental setup used in rule-based target differentiation.




(a) plane

(b) corner

(c) edge

(d) cylinder

Figure 3.2: Intensity-versus-scan-angle characteristics for various targets along the line-of-sight of the experimental setup.

Notice that the preceding (and following) rules are designed to be independent of those features of the scans which vary with range and azimuth, so as to enable position-invariant recognition of the targets. In addition, the proposed method has the advantage that it does not require storage of any reference scans since the information necessary to differentiate the targets are completely embodied in the decision rules.

If the target is not a plane either, we next check whether it is an edge or a cylinder. The intensity patterns for the edge and the cylinder are given in Figures 3.2(c) and (d). They have shapes similar to those of planar targets, but the intersection points of the intensity patterns differ significantly from those of planar targets. In the differentiation between edges and cylinders, we employ the ratio of the intensity value at the intersection of the two scans corresponding to the two sensors, to the maximum intensity value of the scans. (Because the maximum intensity values of the right and left infrared scans are very close, the maximum intensity reading of either infrared sensor or their average can be used in this computation.) This ratio is compared with an empirically determined reference value to determine whether the target is an edge or a cylinder. If the ratio is greater than the reference value, the target is an edge; otherwise, it is a cylinder. (In our experiments, the reference value was 0.65.) If the scan patterns from the two sensors do not intersect, the algorithm cannot distinguish between a cylinder and an edge. However, this never occurred in our experiments. The azimuth estimate of edges and cylinders is also obtained by averaging the angular locations of the maxima of the two scans. Having determined the target type and estimated its azimuth, its range can also be estimated by using linear interpolation between the central values of the individual intensity scans given in Figure 3.2.

The rule-based method is flexible in the sense that by adjusting the threshold parameters of the rules, it is possible to vary the acceptance criterion from tight to loose. If the threshold parameter for the plane is chosen small, the acceptance criterion will be tightened and a greater number of unidentified targets will be produced. If the threshold parameter for the plane is chosen large, the acceptance criterion will be relaxed and targets with greater deviations in geometry (such as 85° corner) or surface properties (such as surface wrapped with rough material)

will be accepted in the same class as the nominal targets.

## 3.2 Experimental Verification

Using the experimental setup described above, the algorithm presented in the previous section is used to differentiate and estimate the position of a plane, a 90° corner, a 90° edge, and a cylinder of radius 4.8 cm.

Based on the results for 160 experimental test scans (from 40 different locations for each target), the target confusion matrix shown in Table 3.1, which contains information about the actual and detected targets, is obtained. The average accuracy over all target types can be found by summing the correct decisions given along the diagonal of the confusion matrix and dividing this sum by the total number of test scans (160), resulting in an average accuracy of 91.3% over all target types. Targets are localized within absolute average range and azimuth errors of 0.55 cm and 1.03°, respectively. The errors have been calculated by averaging the absolute differences between the estimated ranges and azimuths and the actual ranges and azimuths read off from the millimetric grid paper covering the floor of the experimental setup.

The percentage accuracy and confusion rates are presented in Table 3.2. The second column of the table gives the percentage accuracy of correct differentiation of the target and the third column gives the percentage of cases when a certain

Table 3.1: Confusion matrix (P: plane, C: corner, E: edge, CY: cylinder).

| target | differentiation result | | | | total |
|---|---|---|---|---|---|
| | P | C | E | CY | |
| P | 36 | – | 4 | – | 40 |
| C | – | 40 | – | – | 40 |
| E | 4 | – | 33 | 3 | 40 |
| CY | 3 | – | – | 37 | 40 |
| total | 43 | 40 | 37 | 40 | 160 |

Table 3.2: Performance parameters of the algorithm.

| actual target | correct diff. rate (%) | diff. error I (%) | diff. error II (%) |
|---|---|---|---|
| P | 90 | 10 | 16.3 |
| C | 100 | 0 | 0 |
| E | 82.5 | 17.5 | 10.8 |
| CY | 92.5 | 7.5 | 7.5 |
| overall | 91.25 | 8.75 | 8.65 |

target was mistaken for another. The fourth column gives the total percentage of other target types that were mistaken for a particular target type. For instance, for the planar target $(4+3)/43 = 16.3\%$, meaning that targets other than planes are incorrectly classified as planes with a rate of 16.3%.

Because the intensity pattern of a corner differs significantly from the rest of the targets, the algorithm differentiates corners accurately with a rate of 100%. A target is never classified as a corner if it is actually not a corner. Edges and cylinders are the most difficult targets to differentiate.

By designing the decision rules so that they do not depend on those features of the scans which vary with range and azimuth, an average correct target differentiation rate of 91.3% over all target types is achieved and targets are localized within average absolute range and azimuth errors of 0.55 cm and 1.03°, respectively. The proposed method has the advantage that it does not require storage of any reference scans since the information necessary to differentiate the targets are completely embodied in the decision rules. The method also exhibits considerable robustness to deviations in geometry or surface properties of the targets since the rule-based approach emphasizes structural features rather than the exact functional forms of the scans. The major drawback of the present method, as with all such rule-based methods, is that the rules are specific to the set of objects and must be modified for a different set of objects. Nevertheless, the rules we propose are of considerable practical value since the set of objects considered is an important set consisting of the most commonly encountered features in typical

indoor environments and therefore deserves a custom set of rules. (Differentiating this set of objects has long been the subject of investigations involving ultrasonic sensors [1–6].)

We demonstrated differentiation of four basic target types having similar surface properties. Broadly speaking, the major effect of different materials and textures is to change the reflectivity coefficients of the objects. This in turn will primarily have the effect of modifying the amplitudes of the scans with relatively less effect on their structural forms. Therefore, the same general set of rules can be applied with relatively minor modifications or merely adjustments of the parameters.

In the next chapter, we provide template-based differentiation and localization algorithm and extensively investigate the limits of the proposed approach through experimental studies.

# Chapter 4

# TEMPLATE-BASED DIFFERENTIATION

In this chapter, we deal with the problem of differentiating and localizing targets whose geometry and/or surface properties both vary, using a template based approach. This chapter is organized as follows: In Section 4.1, the differentiation and localization of wooden targets of different geometries is proposed [57, 66]. Surface differentiation and localization is explained in Section 4.2 [58, 66]. Section 4.3 deals with simultaneous differentiation and localization of targets whose geometry and surface properties both vary, generalizing and unifying the results of Sections 4.1 and 4.2 [59].

## 4.1 Geometry Differentiation and Localization

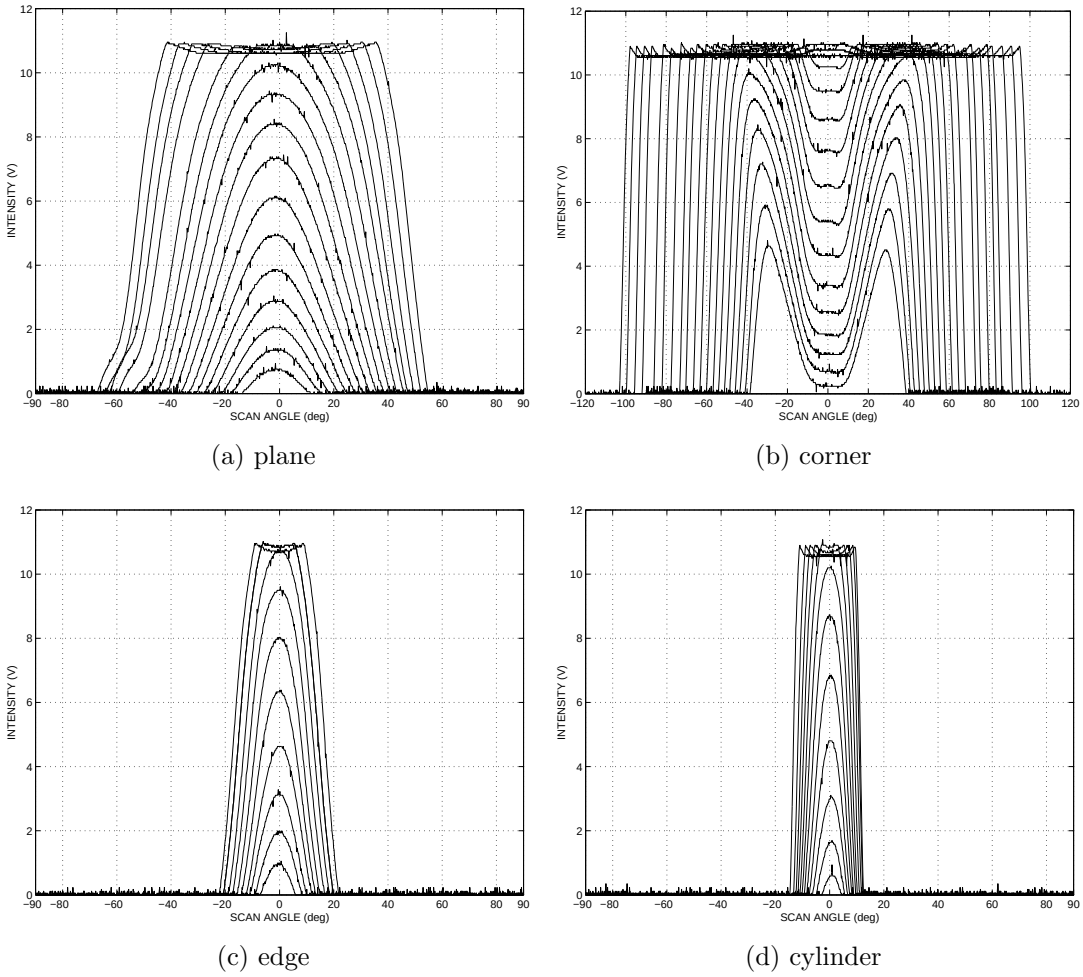
The targets employed in this study are plane, 90° corner, 90° edge, and a cylinder of radius 4.8 cm, whose cross-sections are given in Figure 2.12. They are made of wood, each with a height of 120 cm. Our method is based on angularly scanning each target over a certain angular range. The infrared sensor is mounted on a 12 inch rotary table (Figure 2.11) to obtain angular scans from these target primitives. Reference data sets are collected for each target with 2.5 cm distance

increments, ranging from 15 cm to the maximum detectable range of each target, at $\theta = 0°$.

The resulting reference scans for plane, corner, edge, and cylinder are shown in Figures 4.1, respectively. The intensity scans are $\theta$-invariant but not $r$-invariant; changes in $r$ do not result in any simple scaling. As we will see, these scans contain sufficient information to identify and localize the different target types with a good degree of accuracy. Figure 4.1(b) shows the distinctive double-humped scan pattern for the corner target (this double-humped pattern can be interpreted by thinking of the corner in terms of its two orthogonal constituent planes). The greatest difficulty is encountered in differentiating cylinders and edges which have the most similar intensity patterns. Notice that the return signal intensities saturate at an intensity corresponding to 10.7 V output voltage.

We now describe how to determine the target type and position of an arbitrarily located target whose intensity scan has been observed. First, we check whether the observed scan $I(\alpha)$ exhibits saturation or not. This situation is treated separately as will be explained later in Section 4.1.3.

We start by determining the target type. Unfortunately, direct comparison with the corresponding curves in Figure 4.1 is not possible since we do not yet know the distance of the target, and comparing with all the curves at all distances would be computationally very expensive. Therefore, we exploit the fact that the successive curves in Figure 4.1 exhibit a monotonic dependence on distance. Furthermore, when an observed scan is compared to the several successive curves in any part of Figure 4.1, the two measures of difference between them described in Sections 4.1.1 and 4.1.2 also exhibit a monotonic fall and rise around a single minimum. Therefore, we are assured that we will not be settling at a suboptimal point if we compare the observed scan not with all scans at all distances but only with the four scans (one for each target type) whose central intensities are closest to that of the observed scan. Therefore, for unsaturated scans, only four comparisons need to be made. This remains the case even if the 2.5 cm increments are reduced to smaller values. This has the advantage that the accuracy of the system can be increased without increasing the cost of computation (although a




(a) plane (b) corner

(c) edge (d) cylinder

Figure 4.1: Intensity scans of the four targets at different distances.

greater number of scans do have to be stored). As a test, we also ran a version of the method where *eight* comparisons were made using the scans with the nearest central intensities both above *and* below the observed central intensity, and also using *all* of the scans shown in Figure 4.1. These computationally more expensive approaches (the latter one exceedingly more so) did not improve the results over those of comparison with only four scans. In fact, in the matched filtering case discussed in Section 4.1.2, the results are even somewhat better when four scans are used, due to the fact that this systematic elimination of a *priori* suboptimal scans eliminates the small possibility that they will mistakenly be chosen as the best matching scan due to noise and other errors.

Two alternative approaches are employed in performing the four comparisons. These are discussed below in the following two subsections.

### 4.1.1 Least-Squares Approach

First, we estimate the angular position $\theta$ of the target as follows: Assuming the observed scan pattern is not saturated, we check if it has two humps or not. If so, it is a corner and we find the angular location of the dip in the middle of the two humps and the corresponding intensity value. If not, we find the angular location of the maximum, denoted $\theta_{\mathrm{MAX}}$, and again the corresponding intensity value. These angular values can be directly taken as estimates of the angular position of the target. Alternatively, the angular position can be estimated by finding the center-of-gravity (COG) of the scan as follows:

$$\theta_{\mathrm{COG}} = \frac{\sum_{i=1}^{n} \alpha_i I(\alpha_i)}{\sum_{i=1}^{n} I(\alpha_i)} \tag{4.1}$$

where $n$ is the number of samples in the angular scan. Ideally, these estimates would be equal, but in practice they differ by a small amount. They would be equal under ideal conditions because the scans are symmetric and peaked at their center of symmetry. Symmetry follows from the symmetry of the data acquisition process and the maximum value being at the center is a consequence of the decrease of reflections with increasing $|\alpha|$. We consider the use of both alternatives when tabulating our results. From now on, we will refer to either estimate as the *center angle* of the scan.

Plots of the intensity at the center angle of each scan in Figure 4.1 as a function of the distance at which that scan was obtained, play an important part in our method. Figure 4.2 shows these plots for the maximum intensity (central dip intensity for corner) case.

In this approach, we compare the intensity scan of the observed target with the four reference scans by computing their least-squares differences after aligning their centers with each other. Since the squared difference is sensitive even to multiplicative factors which are close to unity, we have employed an interpolated

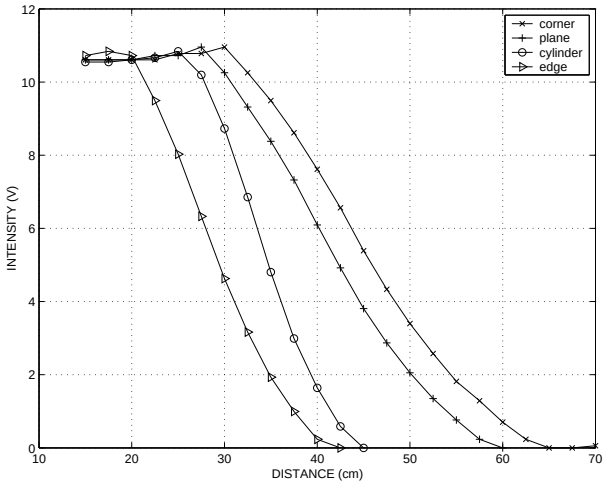



Figure 4.2: Central intensity versus distance curves for the different geometries.

reference scan obtained by linearly interpolating between the two consecutive scans whose central intensities are just above and just below the observed scan. The mean-square difference between the observed scan and the four interpolated scans, one for each possible target type, is computed as follows:

$$\mathcal{E}_j = \frac{1}{n}\sum_{i=1}^{n}[I(\alpha_i - \alpha_{\text{align}}) - I_j(\alpha_i)]^2 \tag{4.2}$$

where $I_j,\ j = 1, 2, 3, 4$ denote the four interpolated reference scans. Here, $\alpha_{\text{align}}$ is the angular shift that is necessary to align the two patterns. The target type resulting in the smallest value of $\mathcal{E}$ is declared as the observed target. Once the target type is determined, the range can be estimated by using linear interpolation on Figure 4.2. We use the set of points associated with the determined geometry type and employ linear interpolation between the points at which reference scans are available to determine a distance estimate from the observed intensity value. For instance, if the geometry is determined to be a corner, and the intensity is observed to be 6 V, we use linear interpolation to estimate the distance as approximately 43.5 cm. Note that, this way, the accuracy of the method is not limited by the 2.5 cm spacing used in collecting the reference scans.

### 4.1.2 Matched Filtering Approach

As an alternative, we have also considered the use of matched filtering [69] to compare the observed and reference scans. The output of the matched filter is the cross-correlation between the observed intensity pattern and the $j$th reference scan normalized by the square root of its total energy:

$$y_j(l) = \frac{\sum_k I(\alpha_k) I_j(\alpha_{k-l})}{\sqrt{\sum_k [I_j(\alpha_k)]^2}} \tag{4.3}$$

where $l = 1, \ldots, 2n-1$ and $j = 1, 2, 3, 4$. The target type corresponding to the maximum cross-correlation peak is declared as the correct target type, and the angular position of the correlation peak directly provides an estimate of the azimuth angle of the target. Then, the distance is estimated by using linear interpolation on Figure 4.2 using the intensity value at the azimuth estimate.

### 4.1.3 Saturated Scans

If saturation is detected in the observed scan, special treatment is necessary. As with other target geometries, a corner scan is considered saturated when its central intensity enters the saturation region, not the humps, since it is the former value which is relevant for our method. In the least-squares approach, mean-square difference between the aligned observed scan and *all* the saturated reference scans are computed and the target type with the minimum mean-square difference is chosen. The range estimate of the target is taken as the distance corresponding to the scan resulting in the minimum mean-square difference. Similarly, for the matched filter, correlation between the observed scan and *all* the stored saturated reference scans is computed and the target type resulting in the highest correlation peak is selected. The range estimate is again taken as that of the best matching scan.

It should be noted that, in the saturated case, range estimation accuracy is limited by the 2.5 cm interval at which the reference scans were taken since

interpolation is not possible. In this case, we cannot expect a maximum error better than $\pm 1.25$ cm and an average absolute error better than 0.625 cm. If this accuracy is not satisfactory, it can be improved by reducing the 2.5 cm intervals. We underline that the 2.5 cm interval does not limit the range estimation accuracy in the unsaturated case, where accurate interpolation is possible from Figure 4.2.

In the unsaturated case, the azimuth could be estimated by taking the angular value corresponding to either the maximum value of the intensity curve or its COG. In the case of saturated scans, a single maximum may not be observed but the COG can still be used to reliably estimate the azimuth. Even when the maximum intensity is used for the unsaturated scans, the COG approach is used for the saturated scans.

### 4.1.4 Experimental Verification

In this section, we experimentally verify the proposed method by locating the targets at randomly selected distances $z$ and azimuth angles $\theta$ and collecting a total of 120 test scans. The targets are randomly located at azimuths varying from $-45^\circ$ to $45^\circ$ from 15 cm up to the maximum ranges in Figure 4.1.

The results of least-squares based target differentiation are displayed in Tables 4.1 and 4.2 in the form of target confusion matrices. Table 4.1 gives the results obtained using the maximum (or the central dip for corner) intensity values, and Table 4.2 gives those obtained using the intensity value at the COG of the scans. The average accuracy over all target types can be found by summing the correct decisions given along the diagonal of the confusion matrix and dividing this sum by the total number of test trials (120). The average correct classification rates obtained by using the max/dip and the COG variations of the least-squares approach are 93% and 89%, respectively.

Matched filter differentiation results are presented in Table 4.3. The average accuracy of differentiation over all target types is 97% which is better than that obtained with the least-squares approach. The matched filter correctly classifies

Table 4.1: Confusion matrix: least-squares based classification (max/dip variation).

| target | differentiation result | | | | total |
|---|---|---|---|---|---|
| | P | C | E | CY | |
| P | 29 | – | 1 | – | 30 |
| C | – | 30 | – | – | 30 |
| E | 1 | – | 26 | 3 | 30 |
| CY | 4 | – | – | 26 | 30 |
| total | 34 | 30 | 27 | 29 | 120 |

Table 4.2: Confusion matrix: least-squares based classification (COG variation).

| target | differentiation result | | | | total |
|---|---|---|---|---|---|
| | P | C | E | CY | |
| P | 30 | – | – | – | 30 |
| C | – | 30 | – | – | 30 |
| E | 5 | – | 23 | 2 | 30 |
| CY | 4 | – | 2 | 24 | 30 |
| total | 39 | 30 | 25 | 26 | 120 |

planar targets as well as corners with an accuracy of 100%.

As shown in the tables, corners are always correctly identified regardless of which method is used, due to their distinctive signature. Second best to corners are planes which are also usually correctly identified. Cylinders and edges are the most confused target types as we had expected from the similar nature of their intensity scans. Nearly all misclassified targets are located at far ranges where the return signal intensities are very weak.

The average absolute range and azimuth estimation errors for the different approaches are presented in Table 4.4 over all test targets. As seen in the table, using the max/dip and COG variations of the least-squares approach, the target

Table 4.3: Confusion matrix: matched filter based classification.

| target | differentiation result | | | | total |
|---|---|---|---|---|---|
| | P | C | E | CY | |
| P | 30 | – | – | – | 30 |
| C | – | 30 | – | – | 30 |
| E | – | – | 29 | 1 | 30 |
| CY | – | – | 3 | 27 | 30 |
| total | 30 | 30 | 32 | 28 | 120 |

Table 4.4: Absolute range and azimuth estimation errors over all test targets. (LS: least-squares, MF: matched filter.)

| method | | P | C | E | CY | average error |
|---|---|---|---|---|---|---|
| LS-max/dip | $r$(cm) | 1.0 | 0.7 | 1.1 | 1.8 | 1.2 |
| | $\theta$(deg) | 4.1 | 5.7 | 2.3 | 1.7 | 3.5 |
| LS-COG | $r$(cm) | 0.5 | 0.7 | 4.3 | 1.5 | 1.7 |
| | $\theta$(deg) | 2.9 | 2.8 | 3.0 | 2.4 | 2.8 |
| MF | $r$(cm) | 0.7 | 0.7 | 0.8 | 1.0 | 0.8 |
| | $\theta$(deg) | 1.2 | 1.7 | 1.8 | 1.8 | 1.6 |

ranges are estimated with average absolute range errors of 1.2 and 1.7 cm, respectively. Matched filtering results in an average absolute range error of 0.8 cm which is much better than that obtained with the least-squares approach. The greatest contribution to the range errors comes from targets which are incorrectly differentiated. If we average over only correctly differentiated targets, the average absolute range errors are reduced to 0.6, 0.6, and 0.7 cm for the max/dip and COG variations of least-squares and the matched filter approaches, respectively. Since these numbers are comparable, we may conclude that the superior range accuracy of matched filtering is mostly a consequence of its superior differentiation accuracy.

As for azimuth estimation, matched filtering results in an average absolute

estimation error of 1.6°, which is the best among the approaches compared. Averaging the azimuth errors over only correctly differentiated targets does not result in significant changes. This is due to the fact that azimuth estimation is not dependent on correct differentiation.

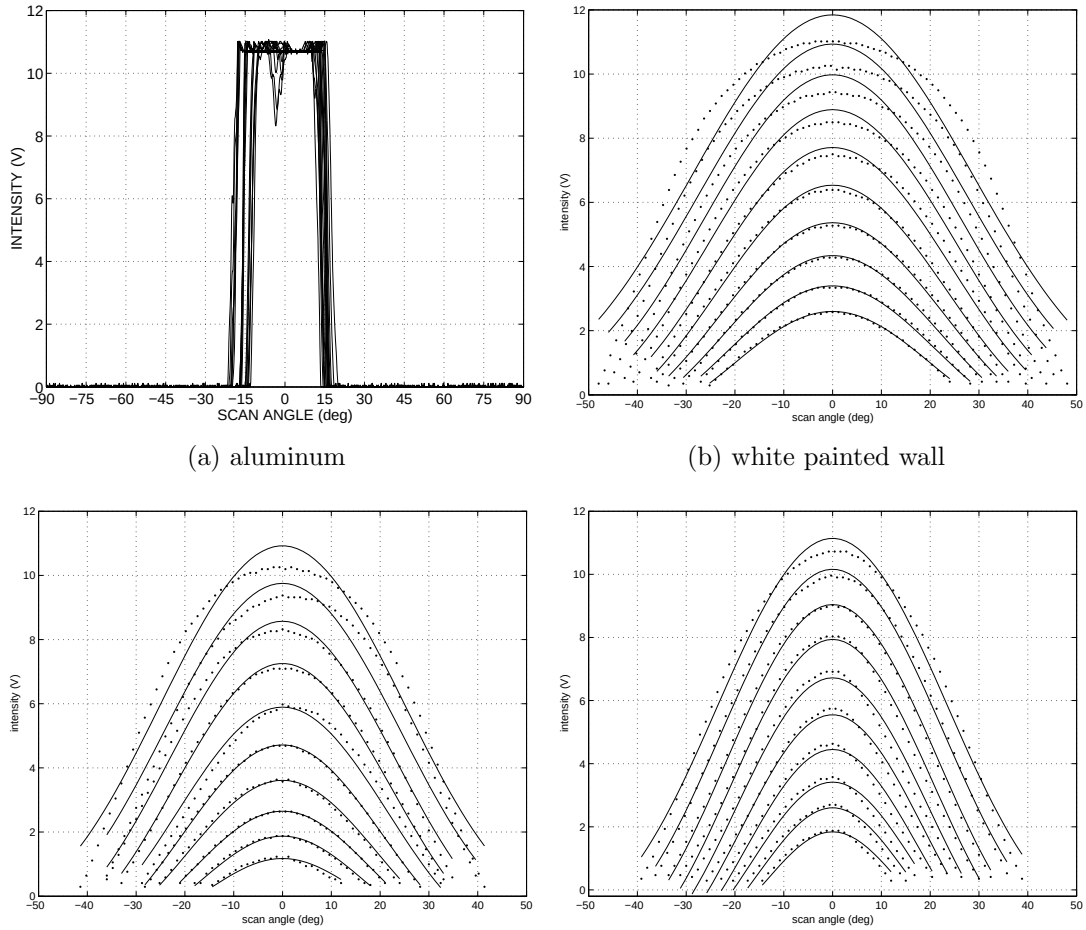
Because of the sharpness of the scans for the cylindrical target around their peaks, azimuth estimation of cylinders is more accurate than that of other targets when the least-squares approach is used. On the other hand, angular localization of corners is less accurate since it is difficult to estimate with good accuracy the exact angular location of the relatively shallow central dip, especially with the max/dip variation of least-squares approach. The COG variation is, on the average, better than the max/dip variation in azimuth estimation due to the fact that COG based calculations average out the noise in the return signal intensities.

The accomplishment of this study is that even though the intensity patterns are highly dependent on target location, and this dependence cannot be represented by a simple relationship, we achieve position-invariant target differentiation. An average correct target differentiation rate of 97% over all target types is achieved and targets are localized within absolute range and azimuth errors of 0.8 cm and 1.6°, respectively. The method we propose is scalable in the sense that the accuracy can be increased by increasing the number of reference scans without increasing the computational cost.

## 4.2 Surface Differentiation and Localization

In this part of the study, we consider differentiation of surfaces having the same planar geometry. The surfaces employed in this study are aluminum, white painted wall, brown craft paper, and Styrofoam packaging material. Our method is based on angularly scanning the surfaces over a certain angular range. Reference data sets are collected for each surface type with 2.5 cm distance increments, ranging from 12.5 cm to 57.5 cm, at $\theta = 0°$.

The resulting reference scans for the four surfaces are shown in Figure 4.3.




(a) aluminum

(b) white painted wall

(c) brown craft paper

(d) Styrofoam packaging material

Figure 4.3: Intensity scans of the four surfaces at different distances.

Notice that the scans are peaked at around $\alpha = 0°$ since both specular and diffuse reflections decrease with increasing $|\alpha|$. The intensity scans are $\theta$-invariant but not $z$-invariant; changes in $z$ do not result in any simple scaling. As we will see, these scans contain sufficient information to identify and localize different surfaces with a good degree of accuracy.

Figure 4.4 shows the maximum intensity of each scan as a function of the distance at which that scan was obtained. Once the type of the surface is determined, the range can be estimated by using Figure 4.4. For instance, if the surface is determined to be a white wall, and the intensity is observed to be 6 V,

we use linear interpolation to estimate the distance as approximately 43.5 cm.

The same methodology as in Section 4.1 is used here.

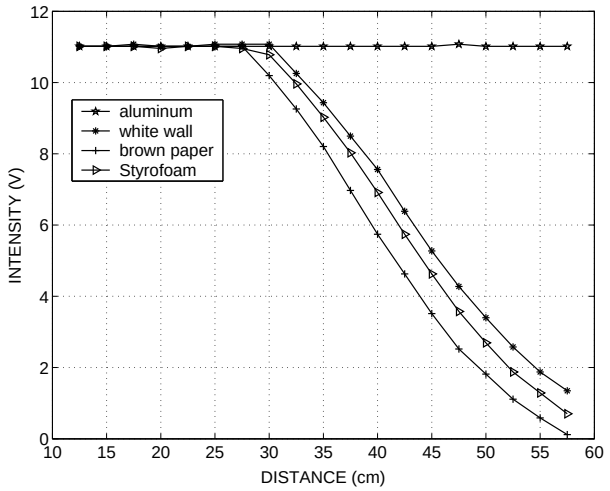



Figure 4.4: Central intensity versus distance curves for the different surfaces.

### 4.2.1 Experimental Verification

In this section, we experimentally verify the proposed method by locating the surfaces at randomly selected distances $z$ and azimuth angles $\theta$ and collecting a total of 100 test scans. The surfaces are randomly located at ranges from 12.5 to 57.5 cm and azimuths from $-45°$ to $45°$.

The results of least-squares based surface differentiation are displayed in Tables 4.5 and 4.6 in the form of surface confusion matrices. Table 4.5 gives the results obtained using the maximum intensity values, and Table 4.6 gives those obtained using the intensity value at the COG of the scans. The average accuracy over all surface types can be found by summing the correct decisions given along the diagonal of the confusion matrix and dividing this sum by the total number of test trials (100). The average correct classification rates obtained by using the maximum intensity and the COG variations of the least-squares approach are 81% and 82%, respectively.

Table 4.5: Confusion matrix: least-squares based recognition (maximum intensity variation). (AL: aluminum, WW: white wall, BR: brown paper, ST: Styrofoam).

| surface | recognition result | | | | total |
|---|---|---|---|---|---|
| | AL | WW | BR | ST | |
| AL | 25 | – | – | – | 25 |
| WW | – | 20 | 3 | 2 | 25 |
| BR | – | 5 | 17 | 3 | 25 |
| ST | – | – | 6 | 19 | 25 |
| total | 25 | 25 | 26 | 24 | 100 |

Table 4.6: Confusion matrix: least-squares based recognition (COG variation).

| surface | recognition result | | | | total |
|---|---|---|---|---|---|
| | AL | WW | BR | ST | |
| AL | 25 | – | – | – | 25 |
| WW | – | 20 | 3 | 2 | 25 |
| BR | – | 4 | 18 | 3 | 25 |
| ST | – | – | 6 | 19 | 25 |
| total | 25 | 24 | 27 | 24 | 100 |

Matched filter differentiation results are presented in Table 4.7. The average accuracy of differentiation over all surfaces is 87%, which is better than that obtained with the least-squares approach. In Section 4.1, where we dealt with the differentiation of targets with different geometries as opposed to different surfaces treated in this section, the least-squares approach resulted in a differentiation accuracy of 93% and 89% and the matched filtering approach resulted in an accuracy of 97%. Based on these results, we conclude that differentiating targets with different surfaces is considerably more difficult than differentiating targets with different geometries.

As shown in the tables, aluminum is always correctly identified regardless of which method is used, due to its distinctive signature. The remaining surfaces are

Table 4.7: Confusion matrix: matched filtering based recognition.

| surface | recognition result | | | | total |
|---|---|---|---|---|---|
| | AL | WW | BR | ST | |
| AL | 25 | – | – | – | 25 |
| WW | – | 21 | 3 | 1 | 25 |
| BR | – | 1 | 21 | 3 | 25 |
| ST | – | – | 5 | 20 | 25 |
| total | 25 | 22 | 29 | 24 | 100 |

comparable in terms of their correct identification percentages. Brown craft paper is the surface most confused with others, especially Styrofoam. Although the intensity scans of these two surfaces do not resemble each other in the unsaturated region, their saturated scans are similar, contributing to the misclassification rate. Nearly all misclassified targets are located at nearby ranges where the return signal intensities are saturated. This means that the misclassification rate can be reduced by increasing the lower limit of the range interval at the cost of reducing the operating range.

The average absolute range and azimuth estimation errors for the different approaches are presented in Table 4.8 over all surface types. As seen in the table, using the maximum intensity and COG variations of the least-squares approach, the target ranges are estimated with average absolute range error of 1.5 cm in both cases. Matched filtering results in an average absolute range error of 1.2 cm which is better than that obtained with the least-squares approach. The greatest contribution to the range errors comes from targets which are incorrectly recognized. If we average over only correctly recognized targets, the average absolute range errors become 1.0, 1.1, and 1.2 cm for the maximum intensity and COG variations of least-squares and the matched filter approaches, respectively. Since these three numbers are relatively closer than the corresponding numbers in Table 4.8, we may conclude that the superior range accuracy of matched filtering is mostly a consequence of its superior differentiation accuracy. The major contribution to range errors comes from saturated scans where linear interpolation from

Table 4.8: Absolute range and azimuth estimation errors over all surfaces.

| method | | AL | WW | BR | ST | average error |
|---|---|---|---|---|---|---|
| LS-max | $r$(cm) | 2.4 | 1.3 | 1.3 | 0.9 | 1.5 |
| | $\theta$(deg) | 0.8 | 1.9 | 1.6 | 0.8 | 1.3 |
| LS-COG | $r$(cm) | 2.4 | 1.3 | 1.3 | 0.9 | 1.5 |
| | $\theta$(deg) | 0.8 | 1.0 | 1.6 | 0.8 | 1.1 |
| MF | $r$(cm) | 1.7 | 1.2 | 1.0 | 0.8 | 1.2 |
| | $\theta$(deg) | 0.8 | 1.1 | 1.6 | 0.7 | 1.0 |

Figure 4.4 cannot be employed to obtain better range estimates. Consequently, surfaces for which saturation occurs over a greater portion of the operating range exhibit greater range estimation errors, with aluminum being the worst.

As for azimuth estimation, matched filtering results in an average absolute estimation error of 1.0°, which is the best among the approaches compared. Averaging the azimuth errors over only correctly differentiated surfaces does not result in significant changes. This is because azimuth estimation is not dependent on correct differentiation. The COG variation is, on the average, better than the maximum intensity variation in azimuth estimation because COG based calculations average out the noise in the return signal intensities.

We have also considered expanding the range of operation of the system. As an example, changing the operating range from [12.5 cm, 57.5 cm] to [5 cm, 60 cm], results in a reduction of the correct differentiation percentage from 87% to 80%. This reduction in performance is mostly a consequence of highly saturated scans and scans with very low intensities, both of which are prone to greater errors.

Light reflected from a surface consists of specular and diffuse components. The specular component is concentrated along the direction where the reflection angle equals the incidence angle, whereas the diffuse component is spread equally

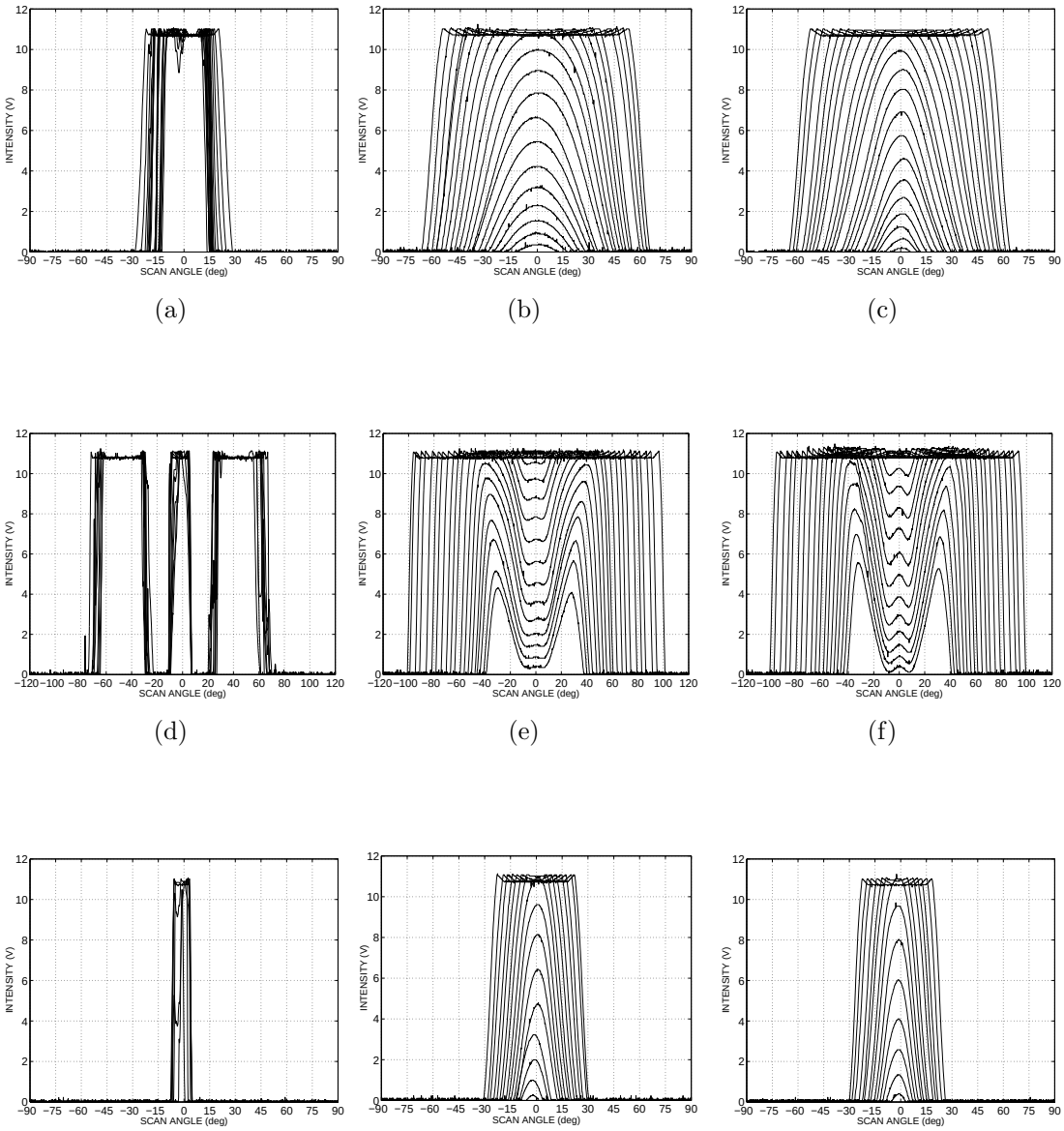
in all directions with a cosine factor. For different types of surfaces, the contribution of these two components and the rate of decrease of intensity with the scan angle $\alpha$ is different. It is this difference which results in a characteristic intensity scan pattern (signature) for each surface, enabling us to distinguish them without knowing their positions. In contrast, a system relying only on reflected energy could not distinguish between a highly reflecting distant object and a less reflecting nearby one. Occasionally, two very distinct surfaces may have intensity scans with very similar dependence on $\alpha$, in which case they cannot be reliably differentiated with the present method.

## 4.3 Geometry and Surface Differentiation and Localization

In this part of the study, we investigate the problem of differentiation and localization targets whose geometry and surface properties both vary, generalizing and unifying the results of Sections 4.1 and 4.2.

The targets employed are plane, 90° corner, and 90° edge each with a height of 120 cm (Figure 2.12). They are covered with aluminum, white cloth, and Styrofoam packaging material. *Reference* data sets are collected for each target with 2.5 cm distance increments, from their nearest to their maximum observable ranges, at $\theta = 0°$.

The resulting reference scans for plane, corner, and edge covered with materials of different surface properties are shown in Figure 4.5. The intensity scans are $\theta$-invariant but not $z$-invariant; changes in $z$ result in variations in both the magnitude and the basewidth of the intensity scans. Scans of corners covered with white cloth and Styrofoam packaging material have a triple-humped pattern (with a much smaller middle hump) corresponding to the two orthogonal constituent planes and their intersection. The intensity scans for corners covered with aluminum [Figure 4.5(d)] have three distinct saturated humps. The same methodology as described in Section 4.2 is used. The same two alternative




(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 4.5: Intensity scans for targets (first row, plane; second row, corner; third row, edge) covered with different surfaces (first column, aluminum; second column, white cloth; third column, Styrofoam) at different distances.

approaches are employed in classifying the observed scans. The only difference is that each observed scan is compared with nine (3 geometries $\times$ 3 surfaces ) reference scans instead of four.

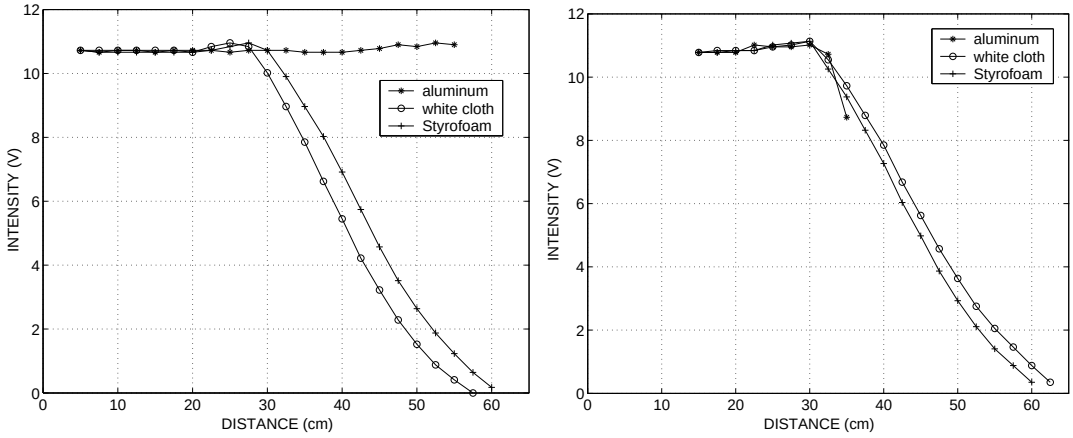
Plots of the intensity at the center angle of each scan in Figure 4.5 as a function of the distance at which that scan was obtained, play an important role in our method. Figure 4.6 shows these plots for the intensity value at the COG of each scan for planes, corners, and edges. Once the geometry and the surface type are determined, the range can be estimated by using linear interpolation on the appropriate curve in Figure 4.6. This way, the accuracy of the method is not limited by the 2.5 cm spacing used in collecting the reference scans.

Table 4.9: Confusion matrix: least-squares based classification (maximum variation) (WC: white cloth).

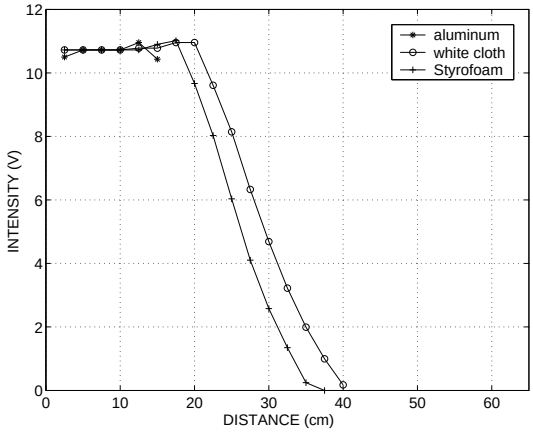
| | | | detected | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | P | | | C | | | E | | |
| | | | AL | WC | ST | AL | WC | ST | AL | WC | ST |
| actual | P | AL | 24 | – | – | – | – | – | – | – | – |
| | | WC | – | 25 | 4 | – | – | – | – | – | – |
| | | ST | – | 9 | 20 | – | – | – | – | – | – |
| | C | AL | – | – | – | 22 | – | – | – | – | – |
| | | WC | – | – | – | – | 10 | 12 | – | – | – |
| | | ST | – | – | – | – | – | 20 | – | – | – |
| | E | AL | – | – | – | – | – | – | 9 | – | 1 |
| | | WC | – | – | – | – | – | – | – | 11 | 9 |
| | | ST | – | – | 1 | – | – | – | – | 8 | 9 |

### 4.3.1 Experimental Verification

In this section, we experimentally verify the proposed method by situating targets at randomly selected distances $z$ and azimuth angles $\theta$ and collecting a total of 194 test scans. The targets are randomly located at azimuth angles varying from $-45°$ to $45°$ from their nearest to their maximum observable ranges in Figure 4.5.




(a) (b)




(c)

Figure 4.6: Central intensity (COG) versus distance curves for different targets: (a) plane; (b) corner; (c) edge.

Table 4.10: Confusion matrix: least-squares based classification (COG variation).

| | | | detected | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | P | | | C | | | E | | |
| | | | AL | WC | ST | AL | WC | ST | AL | WC | ST |
| actual | P | AL | 24 | – | – | – | – | – | – | – | – |
| | | WC | – | 25 | 4 | – | – | – | – | – | – |
| | | ST | – | 9 | 20 | – | – | – | – | – | – |
| | C | AL | – | – | – | 22 | – | – | – | – | – |
| | | WC | – | – | – | – | 13 | 9 | – | – | – |
| | | ST | – | – | – | – | 2 | 18 | – | – | – |
| | E | AL | – | – | 1 | – | – | – | 7 | – | 2 |
| | | WC | – | – | – | – | – | – | – | 14 | 6 |
| | | ST | – | 1 | 1 | – | – | – | – | 10 | 6 |

The results of least-squares based target differentiation are displayed in Tables 4.9 and 4.10 in the form of confusion matrices. Table 4.9 gives the results obtained using the maximum intensity (or the middle-of-two-maxima intensity for corner) values, and Table 4.10 gives those obtained using the intensity value at the COG of the scans. The average accuracy over all target types can be found by summing the correct decisions given along the diagonal of the confusion matrix and dividing this sum by the total number of test trials (194). The same average correct classification rate is achieved by using the maximum and the COG variations of the least-squares approach, which is 77%.

Matched filter differentiation results are presented in Table 4.11. The average accuracy of differentiation over all target types is 80% which is better than that obtained with the least-squares approach.

Planes and corners covered with aluminum are correctly classified with all approaches employed due to their distinctive features. Planar targets of different surface properties are better classified than the others, with a correct differentiation rate of 91% for the matched filtering approach. For corner targets, the highest correct differentiation rate of 83% is achieved with the COG variation

Table 4.11: Confusion matrix: matched filter based classification.

| | | | detected | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | P | | | C | | | E | | |
| | | | AL | WC | ST | AL | WC | ST | AL | WC | ST |
| actual | P | AL | 24 | – | – | – | – | – | – | – | – |
| | | WC | – | 27 | 2 | – | – | – | – | – | – |
| | | ST | – | 5 | 24 | – | – | – | – | – | – |
| | C | AL | – | – | – | 22 | – | – | – | – | – |
| | | WC | – | – | – | – | 14 | 8 | – | – | – |
| | | ST | – | – | – | – | 4 | 16 | – | – | – |
| | E | AL | – | – | – | – | – | – | 9 | 1 | – |
| | | WC | – | – | – | – | – | – | – | 11 | 9 |
| | | ST | – | – | 2 | – | – | – | – | 8 | 8 |

of the least-squares approach. The greatest difficulty is encountered in the differentiation of edges of different surfaces, which have the most similar intensity patterns. The highest correct differentiation rate of 60% for edges is achieved with the maximum intensity variation of the least-squares approach. Taken separately, the geometry and surface type of targets can be correctly classified with rates of 99% and 81%, respectively, which shows that the geometrical properties of the targets are more distinctive than their surface properties, and surface determination is the limiting factor.

The average absolute range and azimuth estimation errors for the different approaches are presented in Table 4.12 for all test targets. As we see in the table, using the maximum and COG variations of the least-squares approach, the target ranges are estimated with average absolute range errors of 1.8 and 1.7 cm, respectively. Matched filtering results in an average absolute range error of 1.5 cm which is better than the least-squares approach. The greatest contribution to the range errors comes from targets which are incorrectly differentiated and/or whose intensity scans are saturated. If we average over only correctly differentiated targets (regardless of whether they lead to saturation), the average absolute range errors are reduced to 1.2, 1.0, and 0.7 cm for the maximum and COG variations of least-squares and the matched filtering approaches, respectively. As for azimuth estimation, the respective average absolute errors for the maximum and COG

Table 4.12: Absolute range and azimuth estimation errors over all test targets.

| method | | P | | | C | | | E | | | average |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AL | WC | ST | AL | WC | ST | AL | WC | ST | error |
| LS-max | $r$(cm) | 2.2 | 2.3 | 1.0 | 2.1 | 0.8 | 0.5 | 2.4 | 1.9 | 2.7 | 1.8 |
| | $\theta$(deg) | 0.9 | 2.3 | 0.8 | 2.4 | 1.7 | 1.3 | 1.1 | 2.0 | 1.7 | 1.6 |
| LS-COG | $r$(cm) | 2.2 | 0.6 | 1.0 | 2.1 | 0.6 | 0.6 | 3.8 | 1.4 | 3.2 | 1.7 |
| | $\theta$(deg) | 0.9 | 1.0 | 0.8 | 2.4 | 1.4 | 1.1 | 1.2 | 2.2 | 2.3 | 1.5 |
| MF | $r$(cm) | 1.7 | 0.5 | 0.7 | 1.5 | 0.6 | 0.6 | 2.2 | 1.7 | 4.2 | 1.5 |
| | $\theta$(deg) | 0.8 | 0.9 | 0.7 | 1.0 | 1.1 | 1.0 | 1.1 | 2.6 | 0.9 | 1.1 |

variations of least-squares and the matched filtering approaches are $1.6°, 1.5°$, and $1.1°$, with matched filtering resulting in the smallest error. When we average over only correctly differentiated targets, these errors are reduced to $1.5°, 1.2°$, and $0.9°$, respectively.

### 4.3.2 Limitations of System Performance

To explore the limitations of the system performance and to assess the robustness of the system, we have also tested the system with targets of either unfamiliar geometry, unfamiliar surface, or both, whose scans are not included in the reference data sets. Therefore, these targets are totally new to the system. First, tests are done for planes, corners, and edges covered with five new surfaces: brown, violet, black, and white paper, and wood. The results of these tests are presented in Tables 4.13–4.15. Planes are classified as planes 100% of the time using both variations of the least-squares method and 99.3% of the time using the matched filtering approach. Corners are classified as corners 100% of the time using any of the three approaches. Edges are correctly classified 89.1% of the time using the maximum variation of the least-squares approach, 88.2% of the time using the COG variation of the least-squares approach, and 87.3% of the time using the matched filtering approach. In these tests, no target type is mistakenly classified as a corner due to the unique characteristics of the corner scans. For the same

Table 4.13: Confusion matrix for planar targets with unfamiliar surface. (WO: wood, VI: violet paper, BL: black paper, WH: white paper.)

| | | | detected | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | P | | | C | | | E | | |
| | | | AL | WC | ST | AL | WC | ST | AL | WC | ST |
| actual | P (LS-max) | WO | – | 16 | 14 | – | – | – | – | – | – |
| | | BR | – | 20 | 10 | – | – | – | – | – | – |
| | | VI | – | 22 | 8 | – | – | – | – | – | – |
| | | BL | – | 24 | 6 | – | – | – | – | – | – |
| | | WH | – | 18 | 11 | – | – | – | – | – | – |
| | P (LS-COG) | WO | – | 15 | 15 | – | – | – | – | – | – |
| | | BR | – | 20 | 10 | – | – | – | – | – | – |
| | | VI | – | 22 | 8 | – | – | – | – | – | – |
| | | BL | – | 24 | 6 | – | – | – | – | – | – |
| | | WH | – | 16 | 13 | – | – | – | – | – | – |
| | P (MF) | WO | – | 19 | 11 | – | – | – | – | – | – |
| | | BR | – | 22 | 8 | – | – | – | – | – | – |
| | | VI | – | 23 | 6 | – | – | – | – | – | 1 |
| | | BL | 1 | 25 | 4 | – | – | – | – | – | – |
| | | WH | – | 18 | 11 | – | – | – | – | – | – |

reason, corners of the preceding five surface types are never classified as planes or edges. The range and azimuth errors are comparable or slightly larger than before (not shown).

We have also tested the system with cylinders, which were not among the three geometries in the original data sets, with the same surface types as used in the reference data sets: aluminum, white cloth, and Styrofoam. The results are given in Table 4.16 and indicate that cylindrical targets are most likely to be classified as edges. In this case, correct surface classification rate drops to 35%. We have also considered cylinders whose surface properties are different than the surface types considered in the reference data sets. These are brown, violet, black, and white paper, and wood. That is, both the geometry and surface type of this target is totally unfamiliar to the system. Again, cylinders are most likely to be classified as edges with Styrofoam surface type (see Table 4.17). In

these two cases, average range estimation error increases to about 9 to 11 cm, but the azimuth error is of the same order of magnitude as before, since our azimuth estimation method is independent of target type.

Table 4.14: Confusion matrix for corner targets with unfamiliar surface.

| | | | detected | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | P | | | C | | | E | | |
| | | | AL | WC | ST | AL | WC | ST | AL | WC | ST |
| actual | C (LS-max) | WO | – | – | – | – | 9 | 13 | – | – | – |
| | | BR | – | – | – | 1 | 3 | 17 | – | – | – |
| | | VI | – | – | – | 1 | 20 | – | – | – | – |
| | | BL | – | – | – | – | 12 | 10 | – | – | – |
| | | WH | – | – | – | – | 12 | 9 | – | – | – |
| | C (LS-COG) | WO | – | – | – | – | 10 | 12 | – | – | – |
| | | BR | – | – | – | 1 | 3 | 17 | – | – | – |
| | | VI | – | – | – | 1 | 2 | 18 | – | – | – |
| | | BL | – | – | – | – | 13 | 9 | – | – | – |
| | | WH | – | – | – | – | 13 | 8 | – | – | – |
| | C (MF) | WO | – | – | – | – | 14 | 8 | – | – | – |
| | | BR | – | – | – | 1 | 4 | 16 | – | – | – |
| | | VI | – | – | – | 1 | 3 | 17 | – | – | – |
| | | BL | – | – | – | – | 13 | 9 | – | – | – |
| | | WH | – | – | – | – | 13 | 8 | – | – | – |

These results indicate that geometrical properties of the targets are more dominant and distinctive compared to their surface properties. When the geometry is familiar but the surface type is not, as in the cases in Tables 4.13–4.15, the correct classification rate of the geometry is very high (about 96% on the average). However, when the surface type is familiar but the geometry is not, the correct classification rate of the surface type is lower (35%), as in Table 4.16.

Among the three approaches, the maximum variation of the least-squares approach is slightly more robust to deviations from targets included in the reference sets.

Table 4.15: Confusion matrix for edge targets with unfamiliar surface.

| | | | detected | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | P | | | C | | | E | | |
| | | | AL | WC | ST | AL | WC | ST | AL | WC | ST |
| actual | E (LS-max) | WO | – | 1 | 5 | – | – | – | – | 9 | 7 |
| | | BR | – | – | 2 | – | – | – | – | 12 | 8 |
| | | VI | – | – | 2 | – | – | – | – | 10 | 8 |
| | | BL | – | – | – | – | – | – | – | 14 | 9 |
| | | WH | – | – | 2 | – | – | – | – | 12 | 9 |
| | E (LS-COG) | WO | – | 2 | 4 | – | – | – | 1 | 11 | 4 |
| | | BR | – | – | – | – | – | – | 1 | 15 | 6 |
| | | VI | – | 1 | 3 | – | – | – | – | 15 | 1 |
| | | BL | – | 1 | – | – | – | – | – | 16 | 6 |
| | | WH | – | 2 | – | – | – | – | – | 13 | 8 |
| | E (MF) | WO | – | – | 6 | – | – | – | – | 12 | 4 |
| | | BR | – | – | 3 | – | – | – | – | 10 | 9 |
| | | VI | – | – | 1 | – | – | – | – | 17 | 2 |
| | | BL | – | – | 2 | – | – | – | – | 15 | 6 |
| | | WH | – | – | 2 | – | – | – | – | 12 | 9 |

### 4.3.3 Effect of Orientation of the Targets

In this section, we will discuss the effect of varying the orientation of the targets from their head-on positions. This constitutes a separate degree-of-freedom than the range and azimuth of the targets. Varying the orientation for planes does not make any difference since a complete scan is acquired. The acquired scan will still be that of a plane, with its peak shifted to the azimuthal value which corresponds to the direction where the sensor line-of-sight is perpendicular to the plane. In other words, varying the orientation of planes does not lead to any deterioration in performance since such planes are already included in the reference set. Variation of orientation is not an issue for cylinders to begin with, since they are rotation invariant.

Change of orientation will make a difference when the target geometry is a corner or an edge, leading to scans not existing in the reference set. Unlike with

Table 4.16: Confusion matrix for cylindrical targets with familiar surface.

| | | | d e t e c t e d | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | P | | | C | | | E | | |
| | | | AL | WC | ST | AL | WC | ST | AL | WC | ST |
| a c t u a l | CY (LS- max) | AL | – | – | – | – | – | – | 1 | – | 12 |
| | | WC | 7 | – | 1 | – | – | – | – | 5 | 12 |
| | | ST | 4 | – | – | – | – | – | 1 | 4 | 16 |
| | CY (LS-COG) | AL | – | – | – | – | – | – | – | – | 13 |
| | | WC | 7 | 1 | – | – | – | – | – | 4 | 13 |
| | | ST | 4 | 1 | 1 | – | – | – | – | 5 | 14 |
| | CY (MF) | AL | – | – | – | – | – | – | 1 | – | 12 |
| | | WC | 8 | – | 2 | – | – | – | – | 2 | 13 |
| | | ST | 5 | – | 1 | – | – | – | – | 5 | 14 |

the case of planes and cylinders, varying the orientation of corners and edges leads to asymmetric scans. If the scan is symmetric, it is either a plane or a cylinder, or a corner or an edge with nearly 0° orientation, and the described algorithm can handle it. If the scan is asymmetric, we know that the target is either a corner or an edge with nonzero orientation. While it is possible to deal with this case by extending the reference set to include targets with nonzero orientation, the introduction of a simple rule allows us to handle such cases with only minor modification of the already presented algorithm. We can determine whether the asymmetric scan comes from a corner or an edge by checking whether or not it has two humps. Thus, even with arbitrary orientations, the target geometry can be determined. Furthermore, we observe that variations in orientation have very little effect on the central intensity of the asymmetric scans (see Figure 4.7 for some examples). This means that the central intensity value can be used for determining the distance in the same manner as before by using linear interpolation on the central intensity versus distance curves for a particular target.

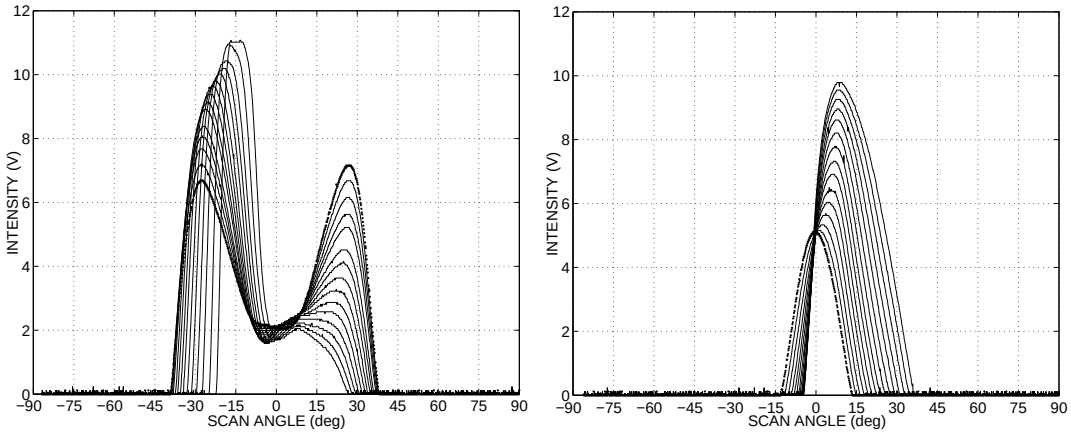
To summarize, with the above observations and minor modifications to the algorithm, the same geometry and surface recognition and position estimation objectives can be achieved even when the targets do not have 0° orientations.

Table 4.17: Confusion matrix for cylindrical targets with unfamiliar surface.

| | | | detected | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | P | | | C | | | E | | |
| | | | AL | WC | ST | AL | WC | ST | AL | WC | ST |
| actual | CY (LS-max) | WO | 8 | – | – | – | – | – | – | 4 | 13 |
| | | BR | 7 | – | – | – | – | – | – | 5 | 13 |
| | | VI | 7 | 1 | 1 | – | – | – | – | 5 | 12 |
| | | BL | 5 | – | – | – | – | – | – | 3 | 16 |
| | | WH | 8 | – | – | – | – | – | – | 5 | 13 |
| | CY (LS-COG) | WO | 8 | – | – | – | – | – | – | 3 | 14 |
| | | BR | 7 | – | 1 | – | – | – | – | 4 | 13 |
| | | VI | 7 | 2 | 1 | – | – | – | – | 5 | 11 |
| | | BL | 5 | – | – | – | – | – | 1 | 7 | 11 |
| | | WH | 8 | 1 | – | – | – | – | – | 3 | 14 |
| | CY (MF) | WO | 8 | – | – | – | – | – | – | 5 | 12 |
| | | BR | 7 | – | 2 | – | – | – | – | 4 | 12 |
| | | VI | 8 | – | 3 | – | – | – | – | 3 | 12 |
| | | BL | 7 | – | 2 | – | – | – | – | 3 | 12 |
| | | WH | 8 | – | – | – | – | – | – | 5 | 13 |

Note, however, that while this approach enables us to accomplish the desired objectives in an orientation-invariant manner, it does not determine the orientation of the target. If determination of target orientation is also desired, this can be accomplished either by storing corresponding scans in the reference set (increasing storage requirements), or more efficiently by constructing orientation angle versus measure-of-asymmetry plots based on suitable measures of asymmetry (for instance, ratios of characteristics of the left- and right-hand sides of the scans).

In order to demonstrate this, we performed additional experiments with corners and edges. These targets were placed at random orientation angles at randomly selected distances. A total of 100 test scans were collected. Using the orientation-invariant approach already described, 100% correct differentiation and absolute mean range errors of 1.02 and 1.47 cm for corners and edges respectively were achieved.




(a) (b)

Figure 4.7: Intensity scans for a wooden (a) corner at 65 cm, (b) edge at 35 cm for orientations between 0° and 35° with 2.5° increments. The curves with the dotted lines indicate 0° orientation.

We also tested the case where reference scans corresponding to different orientations are acquired. Reference data sets were collected for both targets with 5 cm distance increments at $\theta = 0°$, where the orientation of the targets are varied between $-35°$ to $35°$ with 2.5° increments. A total of 489 reference scans were collected. For each test scan, the best-fitting reference scan was found by matched filtering. This method also resulted in 100% correct differentiation rate. Absolute mean range and orientation errors for corners and edges were 1.13 and 1.26 cm and 4.48° and 5.53°, respectively.

The matched filtering approach in general gave better results both for differentiation and localization. The robustness of the methods was investigated by presenting the system with targets of either unfamiliar geometry, unfamiliar surface type, or both. These targets were not included in the reference sets so they were completely new to the system.

An average correct target differentiation rate of 80% over all target types was achieved and targets were localized within absolute range and azimuth errors of 1.5 cm and 1.1°, respectively. The method we propose is scalable in the sense that the accuracy can be increased by increasing the number of reference scans

without increasing the computational cost. The results reported here represent the outcome of our efforts to explore the limits of what is achievable in terms of identifying information with only a simple emitter-detector pair. Such simple sensors are usually put to much lower information-extracting uses.

We have seen that the geometrical properties of the targets are more distinctive than their surface properties, and surface determination is the limiting factor. In this study, we have demonstrated target differentiation for three target geometries and three different surfaces. Based on the data we have collected and the results of Sections 4.1, 4.2, and 4.3, it seems possible to increase the vocabulary of different geometries, provided they are not too similar. However, the same cannot be said for the number of different surfaces. For a given total number of distinct targets, increasing the number of surfaces and decreasing the number of geometries will, in general, worsen the results. On the other hand, decreasing the number of surfaces and increasing the number of geometries will in general improve the results.

In the next chapter, as an alternative to template-based differentiation, we consider processing the same experimental data using artificial neural networks and provide a comparison of the results.

# Chapter 5

# NEURAL NETWORK-BASED DIFFERENTIATION

In this chapter, we propose artificial neural networks for target differentiation as an alternative to the template-based approach described in Chapter 4.

Artificial neural networks (ANNs) have been widely used in areas such as target detection and classification [70], speech processing [71], system identification [72], control theory [73], medical applications [74], and character recognition [75]. In this chapter, ANNs are employed to identify and resolve parameter relations embedded in the characteristics of infrared intensity scans acquired from target types of different geometry, possibly with different surface properties, for their differentiation in a robust manner. This is done in two stages, where the first stage consists of target geometry determination and the second stage involves determining the surface type of the target.

ANNs consist of an input layer, one or more hidden layers to extract progressively more meaningful features, and a single output layer, each comprised of a number of units called *neurons*. The model of each neuron includes a smooth nonlinearity, which is called the *activation function*. Due to the presence of distributed nonlinearity and a high degree of connectivity, theoretical analysis of ANNs is difficult. These networks are trained to compute the boundaries of

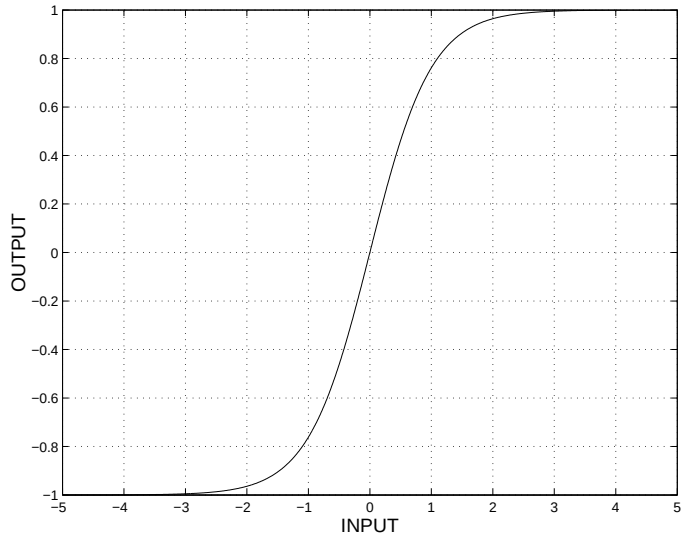



Figure 5.1: Activation function used in the neural networks.

decision regions in the form of connection weights and biases by using training algorithms. Performance of ANNs is affected by the choice of parameters related to the network structure, training algorithm, and input signals, as well as parameter initialization.

This chapter is organized as follows: Section 5.1 introduces the ANN structure and parameters. Geometry and surface determination is described with experimental results in Sections 5.2 and 5.3, respectively.

## 5.1 ANN Structure and Parameters

The ANNs used in this thesis consist of one input, one hidden, and one output layer, with 160, 10, and 3 neurons, respectively. The numbers for the input and hidden layers both include the bias values of 1. The hyperbolic tangent function of the form $\varphi(v) = (1 - e^{-2v})/(1 + e^{-2v})$, illustrated in Figure 5.1, is used as the activation function for all the neurons. The output neurons can take continuous values between –1 and 1 and the decision at the output is made based on a maximum selection scheme. The structure of the ANN is given in Figure 5.2.

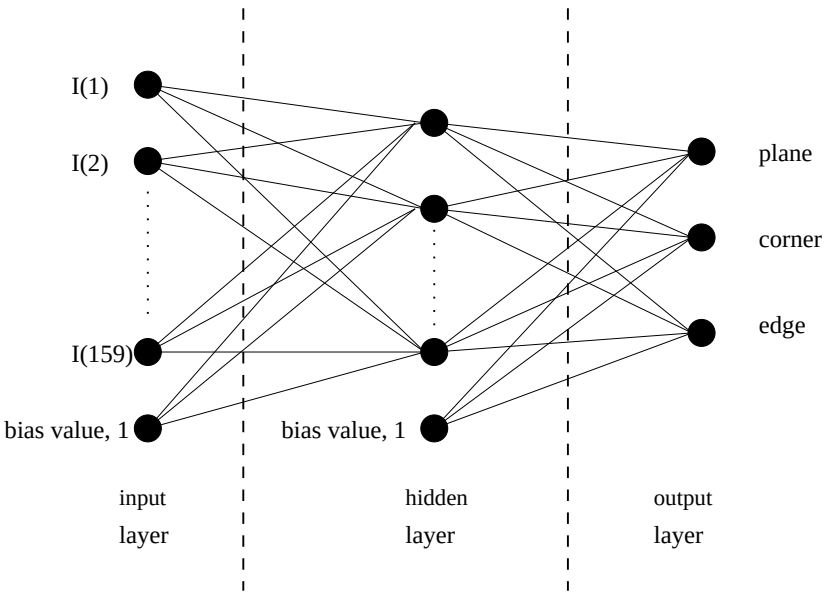



Figure 5.2: Neural network structure used in the study.

It is important to determine the optimal network structure with respect to the correct differentiation rate and network complexity. The infrared scans are provided as input to the ANN after being downsampled by 10 to reduce the complexity of the network (the number of connection weights between the input and hidden layers). This sampling rate is chosen such that the patterns preserve their shapes and no identifying information is lost. Our different trials show that inclusion of more samples of the original scans does not improve the differentiation accuracy. Fully-connected ANNs are trained starting with different initial conditions, different weight factors, and different numbers of neurons in the hidden layer.

Two training algorithms are considered, namely, back-propagation (BP) and Levenberg-Marquardt (LM) algorithms. A set of *training patterns* is presented to the network. With the BP algorithm, the error between the resulting signal at the output and the desired signal is minimized with a gradient-descent procedure. The two adjustment parameters of the algorithm, namely the *learning rate* and the *momentum constant* [76] are chosen as 0.01 and 0.9, respectively.

Since the results of training with BP were not satisfactory, training of ANNs was done by the LM algorithm which is more robust and converges in a shorter

time than BP. When the ANNs are trained with the LM algorithm, a *weight decay factor* is used for regularization. One disadvantage of this algorithm is the high memory requirement, which is not an issue for our relatively small training set. This method is applied in the batch mode, where the network parameters are updated after processing the whole input data. The weight decay factor is chosen as $10^{-4}$. By choosing this factor sufficiently small, one can reduce the average generalization error. If the weight decay factor is chosen too small ($< 10^{-4}$), it will take a long time to converge to the desired accuracy. On the other hand, for greater weight decay factors, the ANN may not converge to the desired accuracy. The learning rate of the LM algorithm is changed adaptively, so the initial value of the learning rate only affects the convergence speed of the network.

Average mean-square error criterion is used for the convergence of ANNs and the acceptable error level is set to a value of $10^{-3}$. The training process stops either when the error criterion is satisfied or the maximum number of epochs (100,000) is exceeded, whichever occurs earlier. The second case occurs very rarely.

After training the fully-connected network, the network structure is further optimized by *pruning* the weights, which involves training the network until it has the minimum number of weights and hidden-layer neurons for a given error tolerance level. Pruning can be done by eliminating weights having the smallest magnitudes, but the resulting ANNs may not be optimal as the weights with smaller magnitudes may be important for the training [77]. In this thesis, Optimal Brain Surgeon technique [55] is implemented for finding the optimal network structure. This technique uses second-order approximation of the criterion function ($J$ in Equation (5.1)) for evaluating the effect of the weights on the training error. $J$ is the sum-squared-error criterion, $J(\mathbf{w}) = \frac{1}{2}\sum_{k=1}^{3}(t_k - z_k)^2$, where $\mathbf{w}$ is the weight vector, $t_k$ and $z_k$ are the desired and actual output values. The network is trained to a local minimum and the weight resulting in the smallest increase in the training error is pruned. The Taylor series expansion of the criterion function around $\mathbf{w}^*$, which results in local minimum training error, is given

by:

$$\delta J = \left(\frac{\delta J}{\delta \boldsymbol{w}}\right)^T \cdot \delta \boldsymbol{w} + \frac{1}{2}\delta \boldsymbol{w}^T \cdot \frac{\delta^2 J}{\delta \mathbf{w}^2} \cdot \delta \boldsymbol{w} + \mathcal{O}\left(\|\delta \mathbf{w}\|^3\right) \tag{5.1}$$

The first term in Equation (5.1) is approximately zero and third and higher-order terms are ignored. The $\frac{\delta^2 J}{\delta \mathbf{w}^2}$ in the second term is called the Hessian matrix ($\mathbf{H}$). Given that one weight is eliminated, the solution that minimizes this function is given as

$$\delta \mathbf{w} = -\frac{w_q}{\left[\mathbf{H}^{-1}\right]_{qq}} \mathbf{H}^{-1} \cdot \mathbf{u}_q \quad \text{and} \quad \mathbf{L}_q = \frac{1}{2} \frac{{w_q}^2}{\left[\mathbf{H}^{-1}\right]_{qq}} \tag{5.2}$$

where $\delta \mathbf{w}$ is optimal weight change, $\mathbf{w}_q$ is the $q$th weight, $\mathbf{u}_q$ is the unit vector along the $q$th direction, $[\mathbf{H}^{-1}]_{qq}$ is the $q$th diagonal element of the inverse Hessian matrix, and $\mathbf{L}_q$ is the training error if weight $q$ is pruned. The weights other than those eliminated are updated using $\mathbf{L}_q$ in Equation (5.2). Optimal Brain Surgeon procedure is outlined below [55]:

1. train a reasonably large network to minimum error
2. compute $\mathbf{H}^{-1}$
3. find the $q$ that gives the smallest saliency and compute $\mathbf{L}_q$
4. if $\mathbf{L}_q$ is less than a preset error bound, eliminate the $q$th weight, proceed to step 5; otherwise proceed to step 6
5. use the $q$ from step 3 to update all weights using $\mathbf{L}_q$
6. If no more weights can be eliminated without large increase in training error, retrain the network

In our case, when 5% of the weights are pruned, the network is retrained within a maximum of 50 iterations. (Retraining can also be done each time one of the weights is pruned. However, this is a very time-consuming process.) At each retraining step, the ANN is tested with the test data and the error and the

corresponding weights are stored. The pruned network resulting in the smallest test error is chosen as the optimal one and is retrained with the weights resulting in the minimum test error, but with no weight decay factor. In the implementation of LM and Optimal Brain Surgeon, the ANN-based system identification toolbox [78, 79] is used.

## 5.2 Differentiation of Geometry Types with ANN

The targets employed in this part of the study are plane, corner, and edge (Figure 2.12), covered with aluminum, white cloth, and Styrofoam packaging material. Reference data sets are collected for each geometry-surface combination with 2.5 cm distance increments, from their nearest to their maximum observable ranges, at $\theta = 0°$. The resulting reference scans are shown in Figure 4.5. Note that these scans are the original scans, not their downsampled versions used as training inputs to the ANN. The training set consists of 147 sample scans, 60 of which correspond to planes, 49 of which correspond to corners, and 38 of which correspond to edges. The number of scans for each geometry is different. This is because the targets have different reflective properties and each target is detectable over a different distance interval determined by its geometry and surface properties. We have chosen to acquire the training scans in a uniformly distributed fashion over the detectable range for each target. Training is done by the LM algorithm. The input weights are initialized randomly. The ANN resulting in the highest correct differentiation rate on the training and test sets has 10 hidden-layer neurons in fully connected form. Initially, Optimal Brain Surgeon is not used for pruning the network so that this ANN may not have the optimal structure.

We test the ANN with infrared data acquired by situating targets at randomly selected distances $r$ and azimuth angles $\theta$ and collecting a total of 194 test scans, 82 of which are from planes, 64 from corners, and 48 from edges. The targets are randomly located at azimuth angles varying from $-45°$ to $45°$ from their nearest

Table 5.1: Confusion matrix for ANN before Optimal Brain Surgeon: results are outside (inside) the parentheses for maximum intensity (COG) based azimuth estimation.

| target | differentiation result | | | total |
|---|---|---|---|---|
| | P | C | E | |
| P | 75(76) | –(–) | 7(6) | 82(82) |
| C | –(–) | 64(64) | –(–) | 64(64) |
| E | –(5) | –(–) | 48(43) | 48(48) |
| total | 75(81) | 64(64) | 55(49) | 194(194) |

to their maximum observable ranges. (Note that the test scans are collected for random target positions and orientations whereas the training set was collected for targets at equally-spaced ranges at $\theta = 0°$.) When a test scan is obtained, first, the azimuth of the target is estimated using the center-of-gravity (COG) and/or the maximum intensity of the scans. The test scans are shifted by the azimuth estimate, then downsampled by 10, and the resulting scan is used as input to the ANN. The differentiation results for the COG case are shown in Table 5.1 in parentheses, where an overall correct differentiation rate of 94.3% is achieved. Corners are always correctly identified and not confused with the other target types due to the special nature of their scans. Planes are confused with edges at six instances out of 82 and similarly, edges are confused with planes in five cases out of 48. Secondly, to observe the effect of the azimuth estimation method, we used the maximum values of the unsaturated intensity scans. The overall correct differentiation rate in this case is 96.4% (given outside the parentheses in Table 5.1), which is better than that obtained using COG, due to the improvement in the classification of edges. Except for seven planar test scans, all planes are correctly differentiated. Six of the seven incorrectly classified planar test targets are covered with aluminum, whose intensity scans are saturated.

At the next step, the ANN is pruned with the Optimal Brain Surgeon technique. The plot of training and test errors with respect to the number of weights left after pruning is shown in Figure 5.3. In this figure, the errors evolve from right to left. The minimum error is obtained on the test set when 263 weights

are used. The eliminated weights are set to zero. As the number of weights is decreased beyond 263, both the training and test errors increase rapidly due to the elimination of too many weights. If 263 weights are kept, the corresponding number of hidden-layer neurons is still 10. Pruned neural network is given in Figure 5.4.

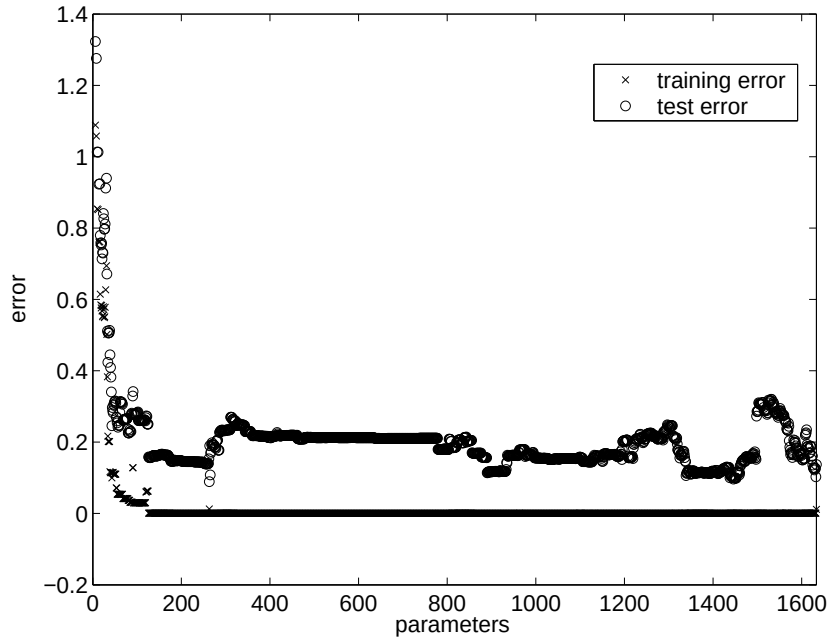



Figure 5.3: Test and training errors while pruning the network with Optimal Brain Surgeon.

Using the weights resulting in the smallest test error, we retrained the network again with the LM algorithm but with zero weight decay factor. The ANN converges in seven iterations to an error of 0.00033. The differentiation results for the optimized network are given in Table 5.2. An overall correct differentiation rate of 99.0% is achieved. Therefore, apart from optimizing the structure of the ANN, pruning the network resulted in improved geometry differentiation.

## 5.3 Differentiation of Surface Types with ANN

In the second stage, we consider differentiating the surface types of the targets assuming their geometries are correctly identified previously. The same network

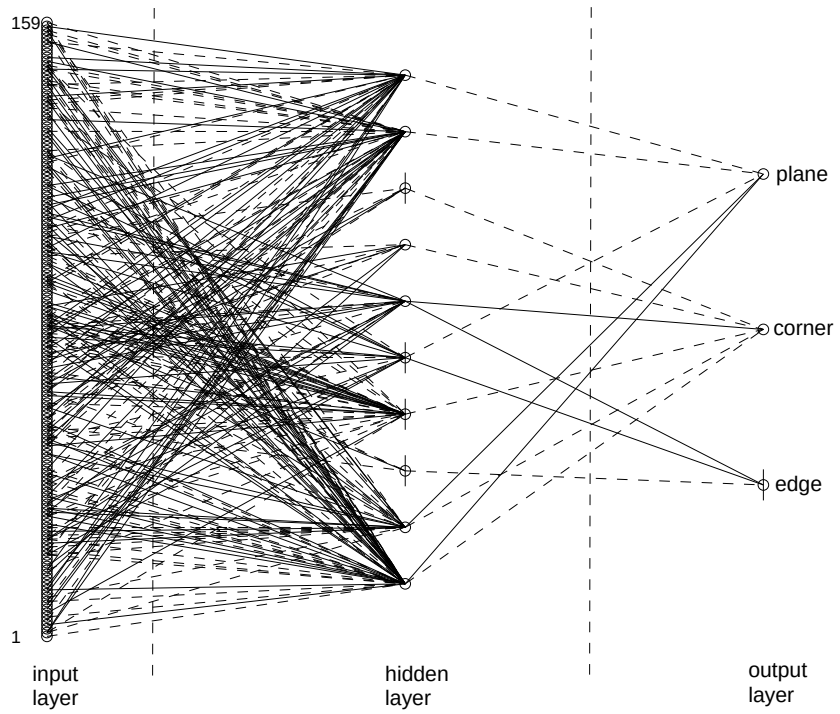



Figure 5.4: Neural network after pruned with Optimal Brain Surgeon.

structure and the same procedure used in geometry differentiation is employed in surface type classification.

For each geometry, all surface types are correctly differentiated in the training set. In Table 5.3, the confusion matrix for the three geometries and surfaces is given. For planes, an average correct differentiation rate of 80.5% is achieved. Planes covered with aluminum are correctly classified with 100% correct differentiation rate. The surface types of the corners are correctly classified with a rate

Table 5.2: Confusion matrix for ANN after Optimal Brain Surgeon.

| target | differentiation result | | | total |
|---|---|---|---|---|
| | P | C | E | |
| P | 80 | – | 2 | 82 |
| C | – | 64 | – | 64 |
| E | – | – | 48 | 48 |
| total | 80 | 64 | 50 | 194 |

Table 5.3: Confusion matrix for three geometries and three surface types.

| | | | d e t e c t e d | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | P | | | C | | | E | | | U |
| | | | AL | WC | ST | AL | WC | ST | AL | WC | ST | |
| a c t u a l | P | AL | 24 | – | – | – | – | – | – | – | – | – |
| | | WC | – | 23 | 6 | – | – | – | – | – | – | – |
| | | ST | – | 9 | 19 | – | – | – | – | – | – | 1 |
| | C | AL | – | – | – | 22 | – | | – | – | – | – |
| | | WC | – | – | – | – | 14 | 8 | – | – | – | – |
| | | ST | – | – | – | – | 1 | 19 | – | – | – | – |
| | E | AL | – | – | – | – | – | – | 8 | – | – | 2 |
| | | WC | – | – | – | – | – | – | – | 19 | – | 1 |
| | | ST | – | – | – | – | – | – | – | 13 | 4 | 1 |

of 85.9%. All corners covered with aluminum are correctly differentiated due to their distinctive features. Worst classification rate (64.6%) is achieved for edges due to their narrower basewidths. Edges covered with Styrofoam are incorrectly classified as edges covered with white cloth with a rate of 72.2%. However, edges covered with white cloth are not confused with Styrofoam packaging material. An overall correct differentiation rate of 78.4% is achieved for all surfaces.

In this chapter, an optimal neural network structure is proposed for improved target classification using a simple infrared sensor. The input signals are the intensity scans obtained from different targets by scanning with a point sensor. The intensity scans are preprocessed by downsampling to decrease the computational complexity of the network. The networks trained with LM are pruned with the Optimal Brain Surgeon technique to obtain the optimal network structure. Pruning also results in improved classification. A modular approach is adopted where first the geometry of the targets is determined, followed by the surface type. Geometry type of the targets is classified with 99% accuracy. Only two planes are incorrectly classified as edges. For the surface type determination, an overall correct differentiation rate of 78.4% is achieved for all surfaces. However, the correct differentiation rates on the test set were worse than expected when compared to our earlier results achieved with the use of techniques other than

neural networks. One of the reasons for this might be the *generalization problem*, where the network memorizes the training set. As the network complexity increases, the network has a tendency to memorize the training set. Complex networks may also result in overtraining where the network tries to learn the noise. On the other hand, simpler network models are not sufficient for the desired tasks. One way to overcome the generalization problem is to divide the test set into training and validation sets. Validation and test sets are constituted by taking every second scan in the original training set. It has been observed that the best differentiation results on the training set do not necessarily occur with the best differentiation rate on the original training set. The main reason for this might be the smaller size of the training set. By dividing the original training set into two smaller subsets, we decrease the representation capability of the training set. Therefore, we tried to increase the number of samples in the training set by adding noise to the scans in the original training set. The noise was added in two ways. First, white Gaussian noise is added to each sample in the scan so that the network does not try to fit to the noisy data. Alternatively, small angular disturbances (fixed for a full scan) are added to the samples so that scans, which might deviate from the $\theta = 0^\circ$ after alignment, can also be correctly differentiated. Another attempt was to reduce the input scan vectors by downsampling the scans. Unfortunately, these trials did not result in any improvement in the correct classification rate.

In the template-based approach in Chapter 4, an average correct classification rate of 80% of both geometry and surface over all target types was achieved. Taken separately, the geometry and surface type of targets were correctly classified with rates of 99% and 81%, respectively. The results are comparable to those achieved in this study.

In the previous chapters, full intensity scans have been used for target differentiation and localization. In the next chapters, we use suitable models for infrared intensity scans and use reflection parameters for target differentiation.

# Chapter 6

# PARAMETRIC DIFFERENTIATION

Our approaches to the differentiation and localization problem in the earlier chapters can be considered as nonparametric where no assumptions about the parametric form of the intensity scans were made. In this chapter, position-invariant surface differentiation is achieved by parametric modeling of the infrared intensity scans rather than using full infrared intensity scans as in the Chapters 4 and 5. The work presented in this chapter was published in [60].

This chapter is organized as follows: In Section 6.1, our parametric modeling of infrared intensity scans is discussed. Section 6.2 provides experimental verification of the proposed approach.

## 6.1 Modeling of Infrared Intensity Scans

Light reflected from a surface depends on the wavelength, the distance, and the properties of the light source (i.e., point or diffuse source), as well as the properties of the surface under consideration such as reflectivity, absorptivity, transmittivity, and orientation [80]. Depending on the surface properties, reflectance can be

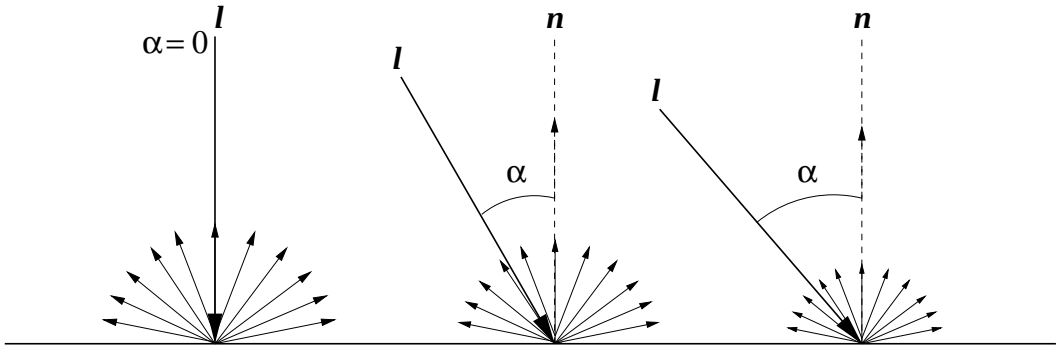



Figure 6.1: Lambertian (diffuse) reflection from an opaque surface. Note how the intensity decreases with increasing $\alpha$ but is of equal magnitude in every direction.

modeled in different ways:

Matte materials can be approximated as ideal Lambertian (diffuse) surfaces which absorb no light and reflect all the incident light equally in all directions such that the intensity of the reflected light is proportional to the cosine of the angle between the incident light and the surface normal [80, 63, 81]. This is known as Lambert's cosine law [82].

When a Lambertian surface is illuminated by a point source of radiance $l_i$, then the radiance reflected from the surface will be

$$l_{s,L} = l_i[k_d(\boldsymbol{l}.\boldsymbol{n})]^d \tag{6.1}$$

where $k_d$ is the coefficient of the diffuse reflection for a given material and $\boldsymbol{l}$ and $\boldsymbol{n}$ are the unit vectors representing the directions of the light source and the surface normal, respectively, as shown in Figure 6.1. If $d > 1$, the diffuse reflection is concentrated on a narrower lobe and resembles specular reflection more. For the case where $d < 1$, the reflected light is more diffused in every direction. Note that in Lambertian reflection, intensity is equally diffused in each direction (therefore, view independent) but the amount of the reflected intensity is dependent on $\cos\alpha$ (Figure 6.1).

In perfect or specular (mirror-like) reflection, the incident light is reflected in the plane defined by the incident light and the surface normal, making an angle with the surface normal which is equal to the incidence angle $\alpha$ [Figure 6.2].

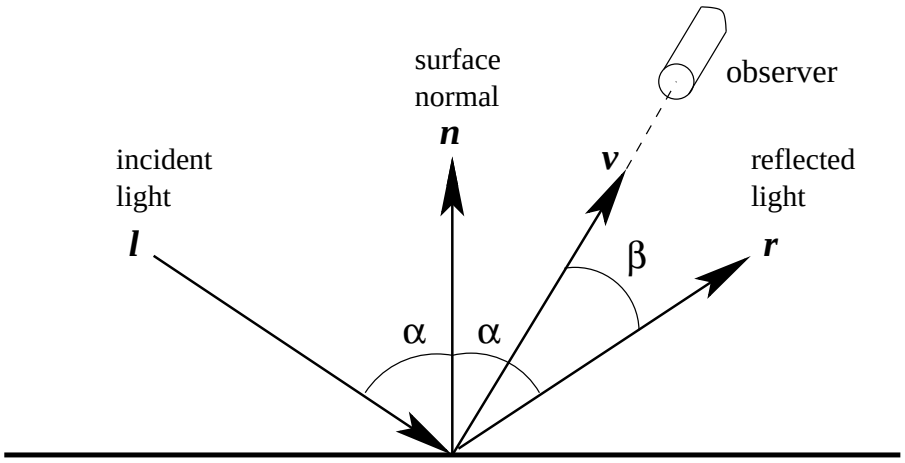



Figure 6.2: Specular reflection from an opaque surface.

For glossy surfaces, the reflected light is approximated as directional diffuse. The radiance reflected from the surface will be

$$l_{s,S} = l_i[k_s(\boldsymbol{r}.\boldsymbol{v})^m] \tag{6.2}$$

where $k_s$ is the coefficient of specular reflection for a given material, $\boldsymbol{r}$ and $\boldsymbol{v}$ are the unit vectors representing the directions of the reflected light and the viewing angle, respectively (Figure 6.2), and $m$ refers to the order of the specular fall-off or shine.

The Phong model [83], which is frequently used in computer graphics applications to represent the intensity of energy reflected from a surface, combines the three types of reflection, which are ambient, diffuse (Lambertian), and specular reflection, in a single formula:

$$l_{s,total} = l_a k_a + l_i[k_d(\boldsymbol{l}.\boldsymbol{n})] + l_i[k_d(\boldsymbol{r}.\boldsymbol{v})^m] \tag{6.3}$$

where $l_{s,total}$ is the total radiance reflected from the surface, $l_a$ and $l_i$ are the ambient and incident radiances on the surface, $k_a, k_d$, and $k_s$ are the coefficients of ambient light, diffuse, and specular reflection for a given material, $\boldsymbol{l}$, $\boldsymbol{n}$, $\boldsymbol{r}$, and $\boldsymbol{v}$ are the unit vectors representing the directions of the light source, the surface normal, the reflected light, and the viewing angle, respectively, as shown in Figure 6.1, and $m$ refers to the order of the specular fall-off or shine as before. The scalar product in the second term of the Phong model equals $\cos\alpha$, where

$\alpha$ is the angle between the vectors $\boldsymbol{l}$ and $\boldsymbol{n}$. Similarly, the scalar product in the last term of the Phong model equals $\cos\beta$ where $\beta$ is the angle between $\boldsymbol{r}$ and $\boldsymbol{v}$. Since the infrared emitter and receiver are situated at approximately the same position, then the angle $\beta$ between the reflected vector $\boldsymbol{r}$ and the viewing vector $\boldsymbol{v}$ is equal to $2\alpha$.

In [41], the simple nonempirical mathematical model represented by Equation 6.3 is used to model reflections from planar surfaces located at a known distance (10 cm) by fitting the reflectance data to the model to improve the accuracy of the range estimates of infrared sensors over a limited range interval (5 to 23 cm). A similar approach with a simplified reflection model is employed in [48], where an infrared sensor-based system can measure distances up to 1 m. The requirement of prior knowledge of the distance to the surface is eliminated in [84, 85] by considering two angular intensity scans taken at two different known distances (10 and 12 cm). The distance error is less than 1 cm over a very limited range interval (10–18 cm) for the reflection coefficients found based on the scans at 10 cm and 12 cm. As the distance increases to the maximum operating range (24 cm), the distance error increases as reported in [84, 85]. For five different surfaces, a correct classification rate of 75% is achieved by considering the invariance property of the sum of the reflection coefficients below a certain range (14 cm) [85]. In the same study, the authors alternatively propose to use the maximum intensity values at a known range for improved surface differentiation, which requires prior knowledge or estimation of the range to the surface. Our approach differs from those in [41, 48] in that it takes distance as a variable and does not require prior knowledge of the distance. Another difference is that those works concentrate mainly on range estimation over a very limited range interval rather than the determination of the surface type, whereas in this thesis, we focus on the determination of the surface type over a broader range interval. When we compare our results with those of [84, 85], we can conclude that the proposed approach is better in terms of the correct differentiation rate and the number of surfaces recognized. Furthermore, in the work done in this thesis, we can simultaneously recognize surfaces and estimate their ranges by relating maximum intensity values to the reflection coefficients in a novel way. The surface

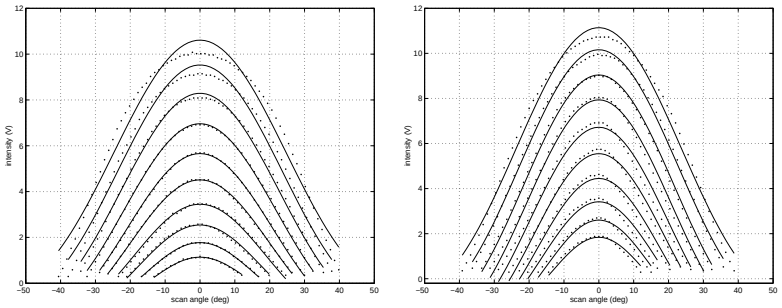
materials considered are unpolished wood, Styrofoam packaging material, white painted matte wall, white and black cloth, and white, brown, and violet paper (not glossy).

Reference intensity scans were collected for each surface type by locating the surfaces between 30 to 52.5 cm with 2.5 cm distance increments, at $\theta = 0°$. The resulting reference scans for the eight surfaces are shown in Figure 6.3 using dotted lines. These intensity scans were modeled by approximating the surfaces as ideal Lambertian surfaces since all of the surface materials involved had matte surfaces. The received return signal intensity is proportional to the detector area and inversely proportional to the square of the distance to the surface and is modeled with three parameters as

$$\mathcal{I} = \frac{C_0 \cos(\alpha C_1)}{\left[\frac{z}{\cos\alpha} + R(\frac{1}{\cos\alpha} - 1)\right]^2} \tag{6.4}$$

which is a modified version of the second term in the model represented by Equation (6.3). In our case, the ambient reflection component, which corresponds to the first term in Equation (6.3), can be neglected with respect to the other terms because the infrared filter, covering the detector window, filters out this term. Furthermore, the second term in Equation (6.3), representing Lambertian reflection, dominates the third term for the matte surface types considered in this study, as further discussed in the following paragraph. In Equation (6.4), the product of the intensity of the emitter, the area of the detector, and the reflection coefficient of the surface is lumped into the constant $C_0$, and $C_1$ is an additional coefficient to compensate for the change in the basewidth of the intensity scans with respect to distance (Figure 6.3). A similar dependence on $C_1$ is used in sensor modeling in [86]. The $z$ is the horizontal distance between the rotary platform and the surface, as shown in Figure 2.11. The denominator of $\mathcal{I}$ is the square of the distance $d$ between the infrared sensor and the surface. From the geometry, $d + R = \frac{z+R}{\cos\alpha}$, from which we obtain $d$ as $\frac{z}{\cos\alpha} + R(\frac{1}{\cos\alpha} - 1)$, where $R$ is the radius of the rotary platform and $\alpha$ is the angle made between the infrared sensor and the horizontal.

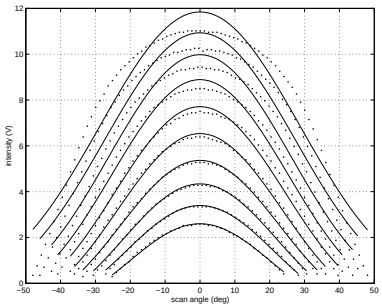
Besides the model represented by Equation (6.4), we have checked the suitability of a number of other models to our experimental data, which were basically



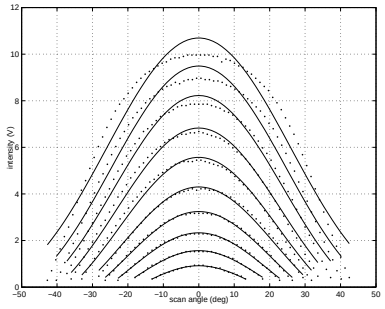

(a) wood



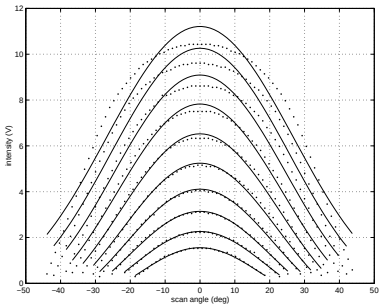

(b) Styrofoam



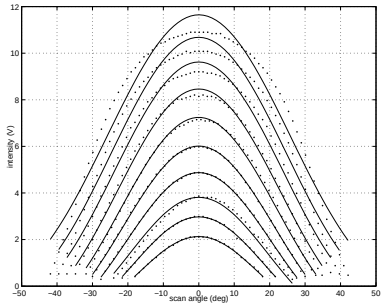

(c) white painted matte wall



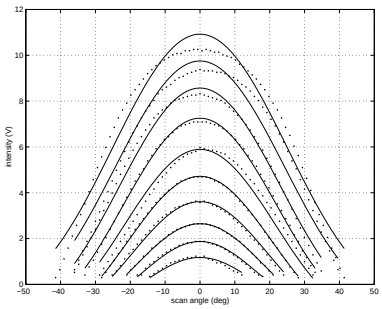

(d) white cloth



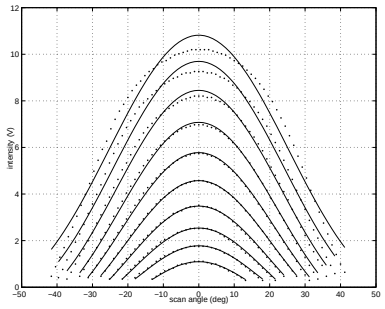

(e) black cloth




(f) white paper




(g) brown paper




(h) violet paper

Figure 6.3: Intensity scans of the eight surfaces collected between 30 to 52.5 cm in 2.5 cm increments. Solid lines indicate the model fit and the dotted lines indicate the experimental data.

different variations of Equation (6.3). The increase in the number of model parameters results in overfitting to the experimental data, whereas simpler models result in larger curve fitting errors. The model represented by Equation (6.4) was the most suitable in the sense that it provided a reasonable trade-off.

Using the model represented by Equation (6.4), parameterized curves were fitted to the reference intensity scans employing a nonlinear least-squares technique based on a model-trust region method provided by MATLAB$^{TM}$ [87]. The resulting curves are shown in Figure 6.3 as solid lines. For the reference scans, $z$ is not taken as a parameter since the distance between the surface and the infrared sensing unit is already known. The initial guesses of the parameters must be made cleverly so that the algorithm does not converge to local minima and curve fitting is achieved in a smaller number of iterations. The initial guess for $C_0$ is made by evaluating $\mathcal{I}$ at $\alpha = 0^\circ$, and corresponds to the product of $\mathcal{I}$ with $z^2$. Similarly, the initial guess for $C_1$ is made by evaluating $C_1$ from Equation (6.4) at a known angle $\alpha$ other than zero, with the initial guess of $C_0$ and the known value of $z$. While curve fitting, $C_0$ value is allowed to vary between $\pm$ 2000 of its initial guess and $C_1$ is restricted to be positive. The variations of $C_0, C_1$, and $z$ with respect to the maximum intensity of the reference scans are shown in Figure 6.4. As the distance $d$ decreases, the maximum intensity increases and $C_0$ first increases then decreases but $C_1$ and $z$ both decrease, as expected from the model represented by Equation (6.4). The model fit is much better for scans with smaller maximum intensities because our model takes only diffuse reflections into account, but the contribution of the specular reflection components around the maximum value of the intensity scans increases as the distance decreases. Hence, the operating range of our system is extended at the expense of the error at nearby ranges.

## 6.2 Experimental Verification

In this section, we experimentally verify the proposed method. In the test process, the surfaces are randomly located at azimuth angles varying from $-45^\circ$ to $45^\circ$, and range values between 30 to 52.5 cm. In the given region, the return signal

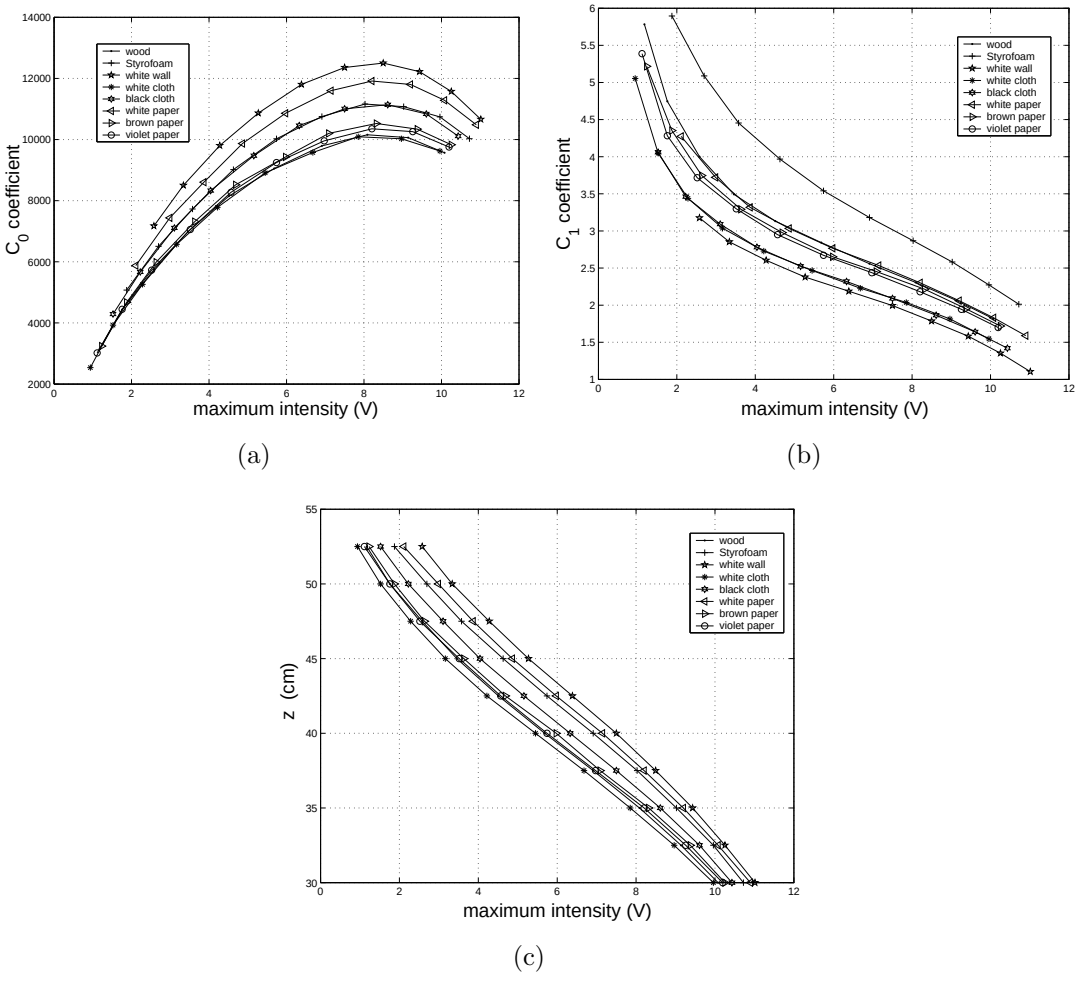



Figure 6.4: Variation of the parameters (a) $C_0$, (b) $C_1$, and (c) $z$ with respect to the maximum intensity of the scan.

intensities do not saturate. In fact, we have experimented with fitting models to the saturated scans so that the operating range of the system is extended to include the saturation regions. However, these trials were not very successful. For unsaturated scans, first, the maximum intensity of the observed intensity scan is found and the angular value where this maximum occurs is taken as the azimuth estimate of the surface. If there are multiple maximum intensity values, the average of the minimum and maximum angular values where the maximum intensity values occur is calculated to find the azimuth estimate of the surface. Then, the observed scan is shifted by the azimuth estimate and the model represented by Equation (6.4) is fitted using a model-trust region based nonlinear least-squares technique [87]. The initial guess for the distance $z$ is found from Figure 6.4(c) by taking the average of the maximum possible and the minimum possible range values corresponding to the maximum value of the recorded intensity scan. (Linear interpolation is used between the data points in the figure.) This results in a maximum absolute range error of approximately 2.5 cm. Therefore, the parameter $z$ is allowed to vary between $\pm 2.5$ cm of its initial guess. Using the initial guess for $z$, the initial guesses for $C_0$ and $C_1$ are made in the same way as already explained for the reference scans. After nonlinear curve fitting to the observed scan, we obtain three parameters $C_0^*, C_1^*$, and $z^*$. In the decision process, the maximum intensity of the observed scan is used and a value of $C_1$ is obtained by linear interpolation between the data points in Figure 6.4(b) for each surface type. In other words, Figure 6.4(b) is used like a look-up table. Surface-type decisions are made based on the absolute difference of $C_1 - C_1^*$ for each surface because of the more distinctive nature of the $C_1$ variation with respect to the maximum intensity. The surface type giving the minimum difference is chosen as the correct one. The decision could have also been made by comparing the parameters with those at the estimated range. However, this would not give better results because of the error and the uncertainty in the range estimates. We have also considered taking different combinations of the differences $C_0 - C_0^*$, $C_1 - C_1^*$, and $z - z^*$ as our error criterion. However, the criterion based on $C_1 - C_1^*$ difference was the most successful.

For a set of six surfaces including Styrofoam packaging material, white painted

matte wall, white or black cloth, and white, brown, and violet paper (also matte), we get a correct differentiation rate of 100% and the surfaces are located with absolute range and azimuth errors of 0.2 cm and 1.1°, respectively. We can increase the number of surfaces differentiated at the expense of a decrease in the correct differentiation rate. For example, if we add wood to our test set keeping either white or black cloth, we get a correct differentiation rate of 86% for seven surfaces (Table 6.1). For these sets of surfaces, absolute range and azimuth

Table 6.1: Confusion matrix: $C_1$-based differentiation (initial range to the surface is estimated using the maximum intensity of the scan).

| surface | differentiation results | | | | | | | total |
|---|---|---|---|---|---|---|---|---|
| | WO | ST | WW | WC(BC) | WP | BR | VI | |
| WO | 4 | – | – | – | 7 | – | 1 | 12 |
| ST | – | 12 | – | – | – | – | – | 12 |
| WW | – | – | 12 | – | – | – | – | 12 |
| WC(BC) | – | – | – | 12 | – | – | – | 12 |
| WP | 4 | – | – | – | 8 | – | – | 12 |
| BR | – | – | – | – | – | 12 | – | 12 |
| VI | – | – | – | – | – | – | 12 | 12 |
| total | 8 | 12 | 12 | 12 | 15 | 12 | 13 | 84 |

errors are 0.6 cm and 1.1°, respectively. Similarly, if we form a set of surfaces excluding wood but keeping both white and black cloth, we achieve a correct differentiation rate of 83% for seven surfaces (Table 6.2) and the surfaces are located with absolute range and azimuth errors of 0.5 cm and 1.1°, respectively. The recognition results for all eight surfaces considered are tabulated in Table 6.3. Over these eight surfaces, an overall correct differentiation rate of 73% is achieved and surfaces are located with absolute range and azimuth errors of 0.8 cm and 1.1°, respectively. Referring to Tables 6.1–6.3, note that the range estimation accuracy improves with increasing correct classification rate, whereas the azimuth estimation accuracy is independent of it, as expected, because of the way it is estimated. In these tables, white and black cloth as well as wood and white paper are the surface pairs most often confused with each other. Thus, the decrease in

Table 6.2: Confusion matrix: $C_1$-based differentiation (initial range to the surface is estimated using the maximum intensity of the scan).

| surface | differentiation results | | | | | | | total |
|---|---|---|---|---|---|---|---|---|
| | ST | WW | WC | BC | WP | BR | VI | |
| ST | 12 | – | – | – | – | – | – | 12 |
| WW | – | 12 | – | – | – | – | – | 12 |
| WC | – | – | 7 | 5 | – | – | – | 12 |
| BC | – | – | 9 | 3 | – | – | – | 12 |
| WP | – | – | – | – | 12 | – | – | 12 |
| BR | – | – | – | – | – | 12 | – | 12 |
| VI | – | – | – | – | – | – | 12 | 12 |
| total | 12 | 12 | 16 | 8 | 12 | 12 | 12 | 84 |

the differentiation rate resulting from adding new surfaces does not represent an overall degradation in differentiation rates across all surface types but is almost totally explained by pairwise confusion of the newly introduced surface with a previously existing one, resulting from the similarity of the $C_1$ parameter of the intensity scans of the two confused surfaces.

To investigate the effect of the initial range estimate of the surface on the differentiation process, we now assume that the distance to the surface is known beforehand. For this case, only the two variables $C_0$ and $C_1$ are taken as parameters. Since the azimuth estimation process is independent of range estimation, for the same set of surfaces, the same azimuth estimation results are obtained. Therefore, they are not repeated here. For the same six surfaces considered as in the previous case (where the initial range to the surface is estimated using the maximum intensity of the scan), the same correct classification rate of 100% is achieved. If we add wood to our test set keeping either white or black cloth, we get a correct differentiation rate of 87% for seven surfaces (Table 6.4). Similarly, if we form a set of surfaces excluding wood but keeping both white and black cloth, we achieve a correct differentiation rate of 88% for seven surfaces (Table 6.5). The differentiation results over all eight surfaces are given in Table 6.6, corresponding to a correct differentiation rate of 78%. When we compare these results with

Table 6.3: Confusion matrix: $C_1$-based differentiation (initial range to the surface is estimated using the maximum intensity of the scan).

| surface | differentiation results | | | | | | | | total |
|---|---|---|---|---|---|---|---|---|---|
| | WO | ST | WW | WC | BC | WP | BR | VI | |
| WO | 4 | – | – | – | – | 7 | – | 1 | 12 |
| ST | – | 12 | – | – | – | – | – | – | 12 |
| WW | – | – | 12 | – | – | – | – | – | 12 |
| WC | – | – | – | 7 | 5 | – | – | – | 12 |
| BC | – | – | – | 9 | 3 | – | – | – | 12 |
| WP | 4 | – | – | – | – | 8 | – | – | 12 |
| BR | – | – | – | – | – | – | 12 | – | 12 |
| VI | – | – | – | – | – | – | – | 12 | 12 |
| total | 8 | 12 | 12 | 16 | 8 | 15 | 12 | 13 | 96 |

those obtained without exact knowledge of the distance to the surface, we can conclude that similar surfaces are confused with each other (wood/white paper and white/black cloth) with smaller confusion rates.

As an alternative, we take as the initial range estimate, the mid-point of the operating range (30 to 52.5 cm), which is 41.25 cm for all surfaces. An overall correct differentiation rate of 65% over eight different surfaces is achieved (Table 6.7), which is worse than the two classification alternatives already considered. The surfaces are located with an absolute range error of 1 cm, which is slightly greater than the absolute range error achieved with the initial range estimate using the maximum intensity of the scan. If we exclude wood and white cloth or wood and black cloth from our test set, we get correct differentiation rates of 93% and 94% for the remaining six surfaces and the surfaces are located with absolute range errors of 0.3 and 0.4 cm, respectively. As azimuth estimation errors are independent of the applied classification techniques, they are not repeated here. Note that for these sets of surfaces, a correct differentiation rate of 100% was achieved using the classification approaches already considered. These high differentiation rates show that even for a maximum initial guess error of 11.25 cm in the range estimates, the proposed approach can recognize a moderate number

Table 6.4: Confusion matrix: $C_1$-based differentiation (range to the surface is known).

| surface | differentiation results | | | | | | | total |
|---|---|---|---|---|---|---|---|---|
| | WO | ST | WW | WC(BC) | WP | BR | VI | |
| WO | 5 | – | – | – | 6 | – | 1 | 12 |
| ST | – | 12 | – | – | – | – | – | 12 |
| WW | – | – | 12 | – | – | – | – | 12 |
| WC(BC) | – | – | – | 12 | – | – | – | 12 |
| WP | 4 | – | – | – | 8 | – | – | 12 |
| BR | – | – | – | – | – | 12 | – | 12 |
| VI | – | – | – | – | – | – | 12 | 12 |
| total | 9 | 12 | 12 | 12 | 14 | 12 | 13 | 84 |

of surfaces with reasonably good accuracy.

The main accomplishment of the parametric approach is that we achieved position-invariant surface differentiation and localization with simple infrared sensors despite the fact that their individual intensity readings are highly dependent on the surface position and properties, and this dependence cannot be represented by a simple analytical relationship. The intensity scan data acquired from a simple low-cost infrared emitter and detector pair was processed and modeled. Different parameterized reflection models were considered and evaluated to find the most suitable model fit to our experimental data, which also best represents and classifies the surfaces under consideration. The proposed approach can differentiate six different surfaces with 100% accuracy. In Chapters 4 and 5, where we considered differentiation and localization of surfaces by employing non-parametric approaches, a maximum correct differentiation rate of 87% over four surfaces was achieved. Comparing this rate with that obtained in this part of the study, we can conclude that the parametric approach is superior to non-parametric ones, in terms of the accuracy, number of surfaces differentiated, and memory requirements, since the non-parametric approaches we considered require the storage of reference scan signals. By parameterizing the intensity scans and

Table 6.5: Confusion matrix: $C_1$-based differentiation (range to the surface is known).

| surface | differentiation results | | | | | | | total |
|---|---|---|---|---|---|---|---|---|
| | ST | WW | WC | BC | WP | BR | VI | |
| ST | 12 | – | – | – | – | – | – | 12 |
| WW | – | 12 | – | – | – | – | – | 12 |
| WC | – | – | 8 | 4 | – | – | – | 12 |
| BC | – | – | 6 | 6 | – | – | – | 12 |
| WP | – | – | – | – | 12 | – | – | 12 |
| BR | – | – | – | – | – | 12 | – | 12 |
| VI | – | – | – | – | – | – | 12 | 12 |
| total | 12 | 12 | 14 | 10 | 12 | 12 | 12 | 84 |

storing only their parameters, we eliminated the need to store complete reference scans. The decrease in the differentiation rate resulting from adding new surfaces in the parametric approach does not represent an overall degradation in differentiation rates across all surface types but is almost totally explained by pairwise confusion of the newly introduced surface with a previously existing one, resulting from the similarity of the $C_1$ parameter of the intensity scans of the two confused surfaces. (Similar decreases in differentiation rate with increasing number of surfaces or objects are also observed with non-parametric template-based approaches.) As an improvement, one can consider using differentiation techniques or learning and/or clustering algorithms that involve more parameters. One possibility is to take a sequential approach: If the estimated $C_1$ parameter of the surface matches more than one surface closely, one can then inspect the other parameters of the surface in sequence. This would be faster than taking all the parameters into account all of the time.

In the next chapter, target properties are determined based on the parametric modeling discussed in this chapter using statistical pattern recognition techniques.

Table 6.6: Confusion matrix: $C_1$-based differentiation (range to the surface is known).

| surface | differentiation results | | | | | | | | total |
|---|---|---|---|---|---|---|---|---|---|
| | WO | ST | WW | WC | BC | WP | BR | VI | |
| WO | 5 | – | – | – | – | 6 | – | 1 | 12 |
| ST | – | 12 | – | – | – | – | – | – | 12 |
| WW | – | – | 12 | – | – | – | – | – | 12 |
| WC | – | – | – | 8 | 4 | – | – | – | 12 |
| BC | – | – | – | 6 | 6 | – | – | – | 12 |
| WP | 4 | – | – | – | – | 8 | – | – | 12 |
| BR | – | – | – | – | – | – | 12 | – | 12 |
| VI | – | – | – | – | – | – | – | 12 | 12 |
| total | 9 | 12 | 12 | 14 | 10 | 14 | 12 | 13 | 96 |

Table 6.7: Confusion matrix: $C_1$-based differentiation (initial range estimate is taken as half of the operating range for all surfaces).

| surface | differentiation results | | | | | | | | total |
|---|---|---|---|---|---|---|---|---|---|
| | WO | ST | WW | WC | BC | WP | BR | VI | |
| WO | 2 | – | – | – | – | 9 | – | 1 | 12 |
| ST | – | 12 | – | – | – | – | – | – | 12 |
| WW | – | – | 9 | 1 | 2 | – | – | – | 12 |
| WC | – | – | – | 7 | 5 | – | – | – | 12 |
| BC | – | – | – | 10 | 2 | – | – | – | 12 |
| WP | 4 | – | – | – | – | 7 | 1 | – | 12 |
| BR | 1 | – | – | – | – | – | 11 | – | 12 |
| VI | – | – | – | – | – | – | – | 12 | 12 |
| total | 7 | 12 | 9 | 18 | 9 | 16 | 12 | 13 | 96 |

# Chapter 7

# DIFFERENTIATION BASED ON STATISTICAL PATTERN RECOGNITION TECHNIQUES

In this chapter, we extend the parametric surface differentiation approach presented in the previous chapter to the differentiation of the geometry of the target types in parameter space, using statistical pattern recognition techniques. Part of this work is published in [61] and it is also submitted as a full paper.

## 7.1 Statistical Pattern Recognition Techniques

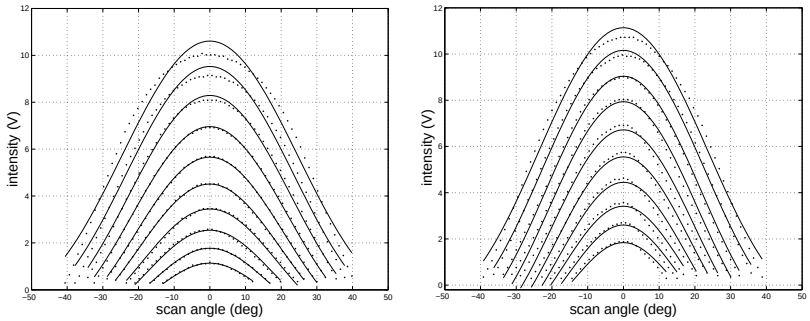
The geometries considered are plane, edge, and cylinder made of unpolished oak wood. The surfaces are either left uncovered (plain wood) or alternatively covered with Styrofoam packaging material, white and black cloth, and white, brown, and violet paper (matte). In the implementation, PRTools [88, 89] is used.

After nonlinear curve fitting to the observed scan as in Chapter 6 (see Figures 7.1, 7.2, and 7.3), we get three parameters $C_0, C_1$, and $z$. We begin by

constructing two alternative feature vector representations based on the parametric representation of the infrared scans. The *feature vector* $\mathbf{x}$ is a $2 \times 1$ column vector comprised of either the $[C_0, I_{\max}]^T$ or the $[C_1, I_{\max}]^T$ pair, illustrated in Figures 7.4 (a) and (b), respectively. Therefore, the dimensionality $d$ of the feature vector representations is 2.

We associate a class $w_i$ with each target type ($i = 1, \ldots, c$). An unknown target is assigned to class $w_i$ if its feature vector $\mathbf{x} = [x_1, \ldots, x_d]^T$ falls in the region $\Omega_i$. A rule which partitions the decision space into regions $\Omega_i, i = 1, \ldots, c$ is called a *decision rule*. Each one of these regions corresponds to a different target type. Boundaries between these regions are called *decision surfaces*. Let $p(w_i)$ be the a *priori* probability of a target belonging to class $w_i$. To classify a target with feature vector $\mathbf{x}$, a *posteriori* probabilities $p(w_i|\mathbf{x})$ are compared and the target is classified into class $w_j$ if $p(w_j|\mathbf{x}) > p(w_i|\mathbf{x}) \quad \forall i \neq j$. This is known as *Bayes minimum error rule*. However, since these a *posteriori* probabilities are rarely known, they need to be estimated. A more convenient formulation of this rule can be obtained by using Bayes' theorem: $p(w_i|\mathbf{x}) = p(\mathbf{x}|w_i)p(w_i)/p(\mathbf{x})$ which results in $p(\mathbf{x}|w_j)p(w_j) > p(\mathbf{x}|w_i)p(w_i) \quad \forall i \neq j \implies \mathbf{x} \in \Omega_j$ where $p(\mathbf{x}|w_i)$ are the class-conditional probability density functions (CCPDFs) which are also unknown and need to be estimated in their turn based on the training set. The training set consists of several sample feature vectors $\mathbf{x}_n, \ n = 1, \ldots, N_i$ which all belong to the same class $w_i$, for a total of $N_1 + N_2 + \ldots + N_c = N$ sample feature vectors. The *test set* is then used to evaluate the performance of the decision rule used. This decision rule can be generalized as $q_j(\mathbf{x}) > q_i(\mathbf{x}) \quad \forall i \neq j \implies \mathbf{x} \in \Omega_j$ where the function $q_i$ is called a *discriminant function*.

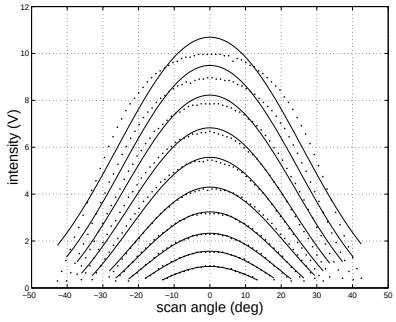
The various statistical techniques for estimating the CCPDFs based on the training set are often categorized as non-parametric and parametric. In non-parametric methods, no assumptions on the parametric form of the CCPDFs are made; however, this requires large training sets. This is because any non-parametric PDF estimate based on a finite sample is biased [90]. In parametric methods, specific models for the CCPDFs are assumed and then the parameters of these models are estimated. These parametric methods can be categorized as normal and non-normal models.



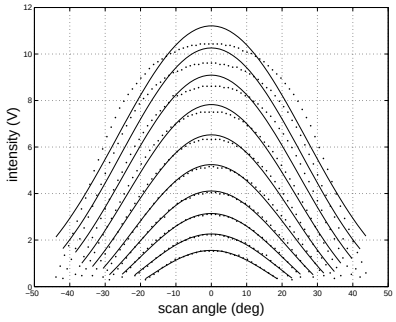

(a) wood



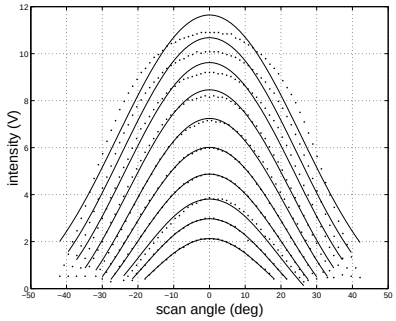

(b) Styrofoam



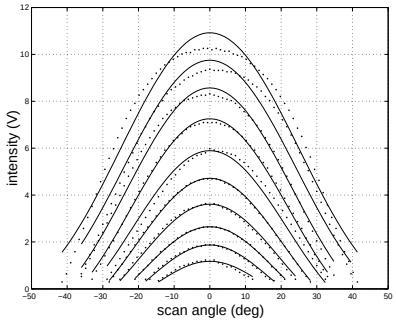

(c) white cloth



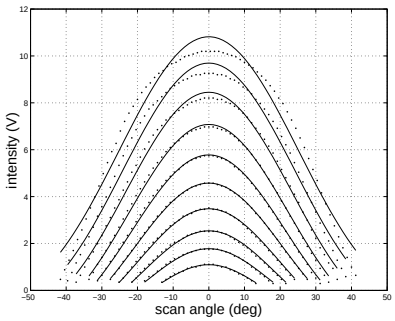

(d) black cloth




(e) white paper

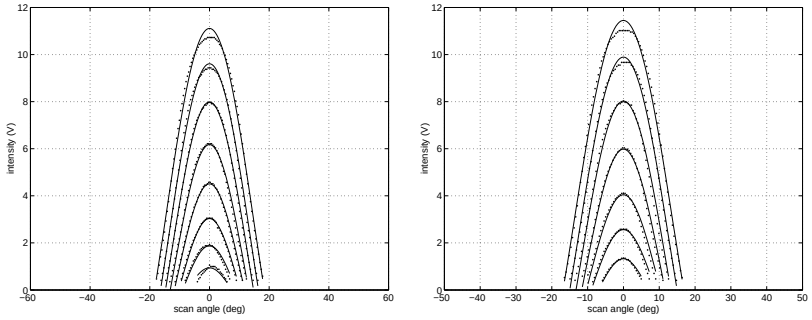



(f) brown paper




(g) violet paper

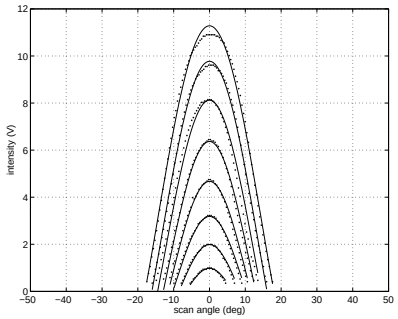
Figure 7.1: Intensity scans of the planes covered with seven planar surfaces collected at different ranges [see Figure 7.4(c)]. Solid lines indicate the model fit and the dotted lines indicate the actual data.



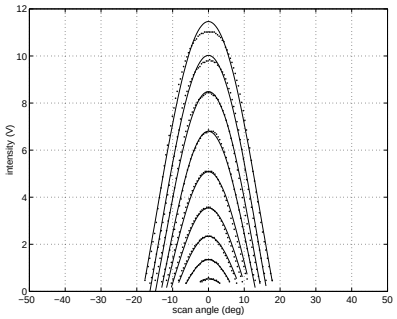

(a) wood



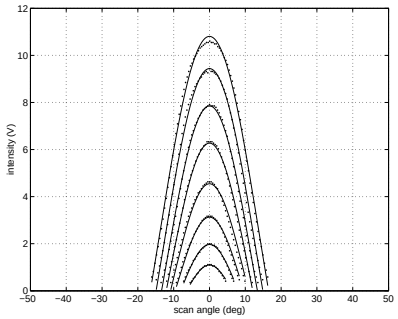

(b) Styrofoam



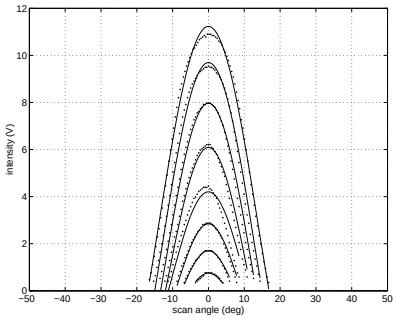

(c) white cloth



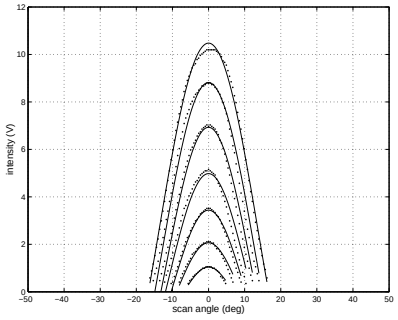

(d) black cloth




(e) white paper

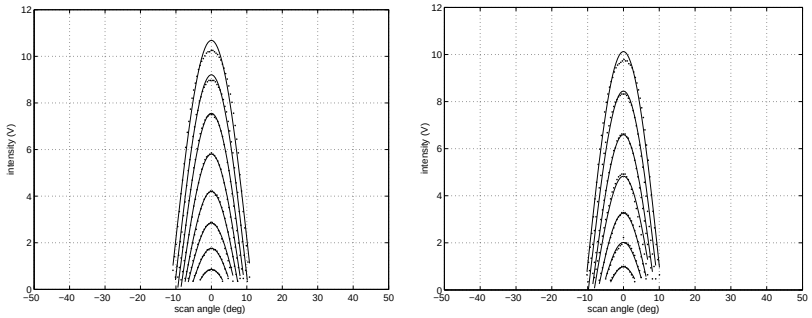



(f) brown paper

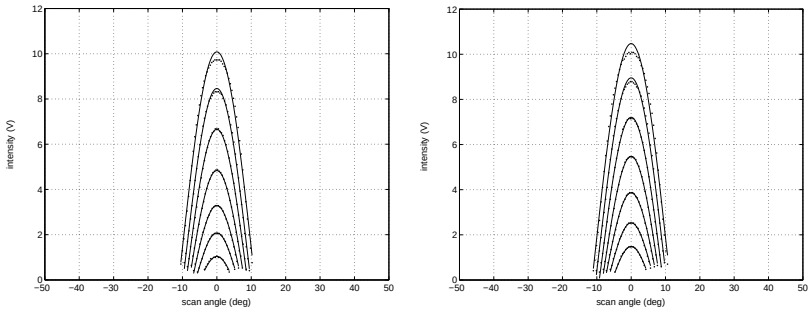



(g) violet paper

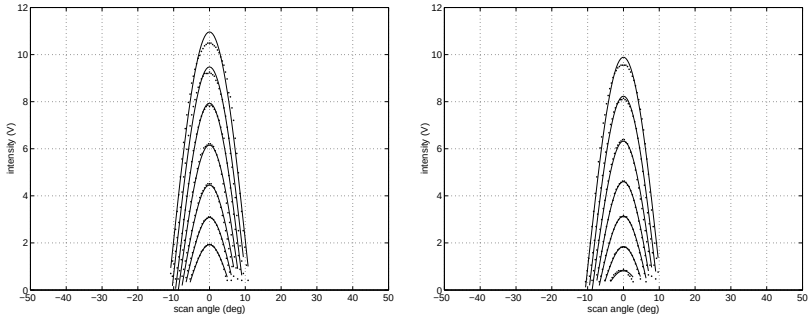
Figure 7.2: Intensity scans of the edges covered seven surfaces collected at different ranges [see Figure 7.4(c)]. Solid lines indicate the model fit and the dotted lines indicate the actual data.



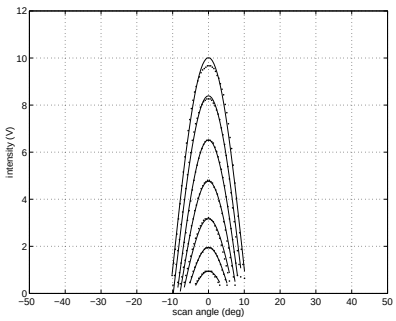

(a) wood (b) Styrofoam




(c) white cloth (d) black cloth




(e) white paper (f) brown paper




(g) violet paper

Figure 7.3: Intensity scans of the cylinders covered with seven surfaces collected at different ranges [see Figure 7.4(c)]. Solid lines indicate the model fit and the dotted lines indicate the actual data.

### 7.1.1 Determination of Geometry

#### 7.1.1.1 Normal Density Based Classifiers

**7.1.1.1.1 Parameterized Density Estimation (PDE):** In this method, the CCPDFs are assumed to be $d$-dimensional normal:

$$p(\mathbf{x}|w_i) = \frac{1}{(2\pi)^{(d/2)}|\mathbf{\Sigma}_i|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T\mathbf{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right], \quad i = 1,\ldots,c \quad (7.1)$$

where the $\boldsymbol{\mu}_i$'s denote the class means, and the $\mathbf{\Sigma}_i$'s denote the class-covariance matrices, both of which must be estimated based on the training set. The most commonly used parameter estimation technique is the maximum likelihood estimator (MLE) [91] which is also used in this study.

In PDE, $d$-dimensional *homoscedastic* and *heteroscedastic* normal models are used for the CCPDFs. In the homoscedastic case, the covariance matrices for all classes are selected equal, usually taken as a weighted (by a *priori* probabilities) average of the individual class-covariance matrices: $\sum_{i=1}^{c} \frac{N_i}{N}\hat{\mathbf{\Sigma}}_i$ [92]. In the heteroscedastic case, they are individually calculated for each class.

In this study, both homoscedastic and heteroscedastic normal models have been implemented to estimate the means and the covariances of the CCPDF for each class (i.e., target type) using the MLE, for each of the two feature vector representations described above. These are the $[C_0, I_{\max}]^T$ and $[C_1, I_{\max}]^T$ feature vectors illustrated in Figure 7.4(a) and (b), respectively.

The training set consists of $N = 175$ data pairs for three classes: $N_1 = 70$

Table 7.1: Confusion matrix: homoscedastic PDE using the $[C_0, I_{\max}]^T$ feature vector. Numbers outside (inside) the parentheses are for the training (test) scans.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 61(–) | –(50) | 9(34) | 70(84) |
| E | –(–) | 49(43) | 6(–) | 55(43) |
| CY | 4(–) | 5(84) | 41(–) | 50(84) |
| total | 65(–) | 54(177) | 56(34) | 175(211) |

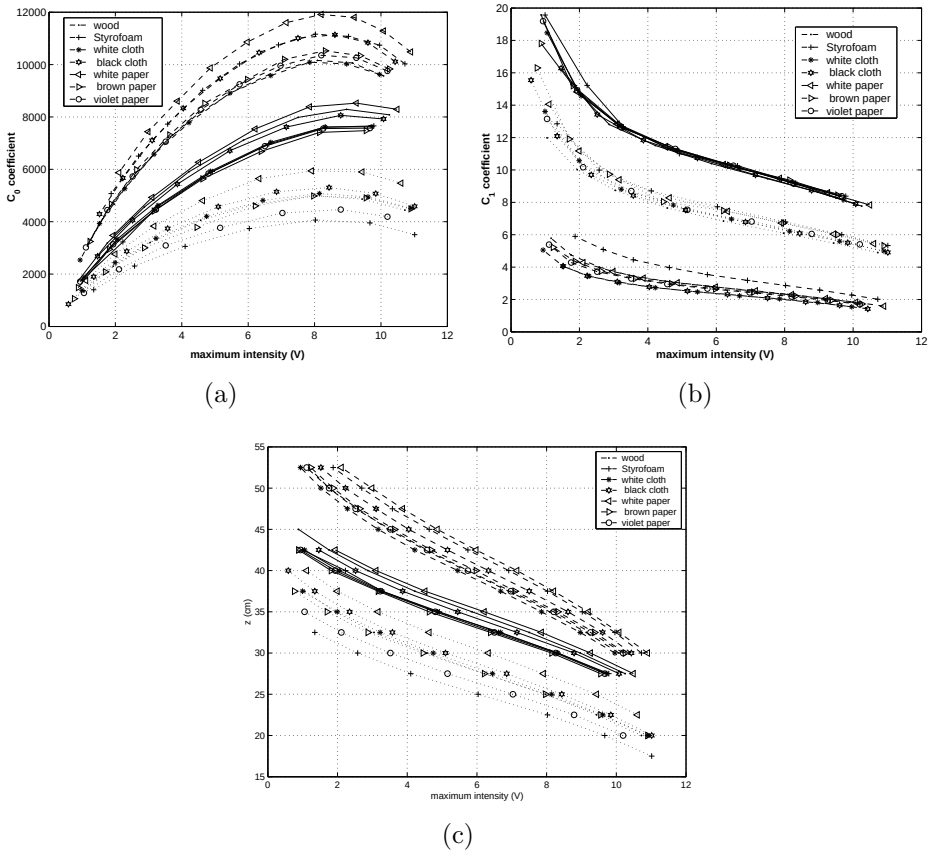



Figure 7.4: Variation of the parameters (a) $C_0$, (b) $C_1$, and (c) $z$ with respect to maximum intensity (dashed, dotted, and solid lines are for planes, edges, and cylinders, respectively).

planes, $N_2 = 55$ edges, and $N_3 = 50$ cylinders. The test set consists of 211 data pairs for three classes: 84 cylinders, 43 edges, and 84 planes. A given test feature vector is classified into the class for which Equation (7.1) is maximum.

Since the feature vector size $d$ is two and the number of classes $c$ is three, three 2-D normal functions are used for classification. The discriminant functions for PDE are plotted on the training set feature vectors $[C_0, I_{\max}]^T$ in Figure 7.5. The classification results are given in Table 7.1 for both the training and test sets for homoscedastic PDE. Overall correct differentiation rates of 86.3% and 20.4% are achieved for the training and test sets, respectively. The main reason for the low differentiation rate on the test set is due to the $[C_0, I_{\max}]^T$ feature vector of the observed intensity scans not being very distinctive. For the heteroscedastic case,

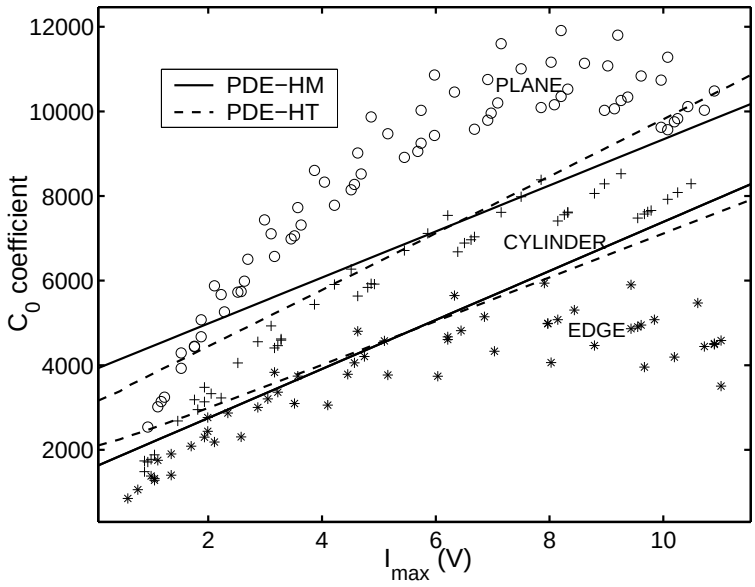


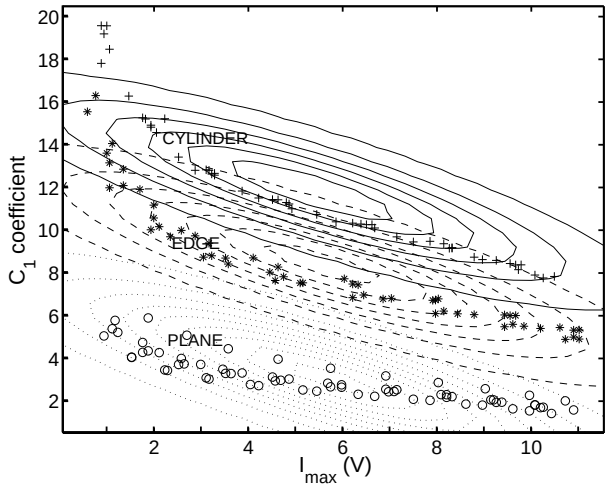

Figure 7.5: Discriminant functions for PDE when the $[C_0, I_{\max}]$ feature vector is used.

the geometry confusion matrix is given in Table 7.2. The differentiation rates for this case are same as the homoscedastic case.

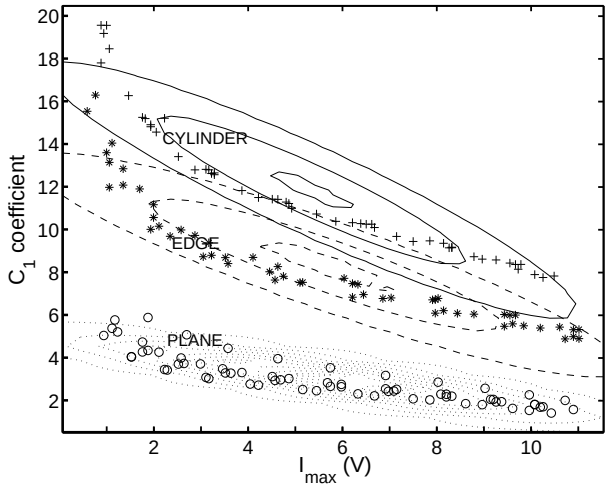
Equal probability contours of the 2-D normal functions are given in Figure 7.6 (a) and (b) for each case when the $[C_1, I_{\max}]^T$ feature vector is used for differentiation. The corresponding discriminant functions are shown in Figure 7.7. From Table 7.3, the correct differentiation rates for homoscedastic PDE are 96.6% and 98.6% for the training and test sets, respectively. For the test data,

Table 7.2: Confusion matrix: heteroscedastic PDE using the $[C_0, I_{\max}]^T$ feature vector.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 60(–) | –(46) | 10(38) | 70(84) |
| E | –(–) | 51(43) | 4(–) | 55(43) |
| CY | 4(–) | 6(84) | 40(–) | 50(84) |
| total | 64(–) | 57(173) | 54(38) | 175(211) |




(a)




(b)

Figure 7.6: 2-D normal contour plots for (a) homoscedastic (b) heteroscedastic PDE when the $[C_1, I_{\max}]^T$ feature vector is used.

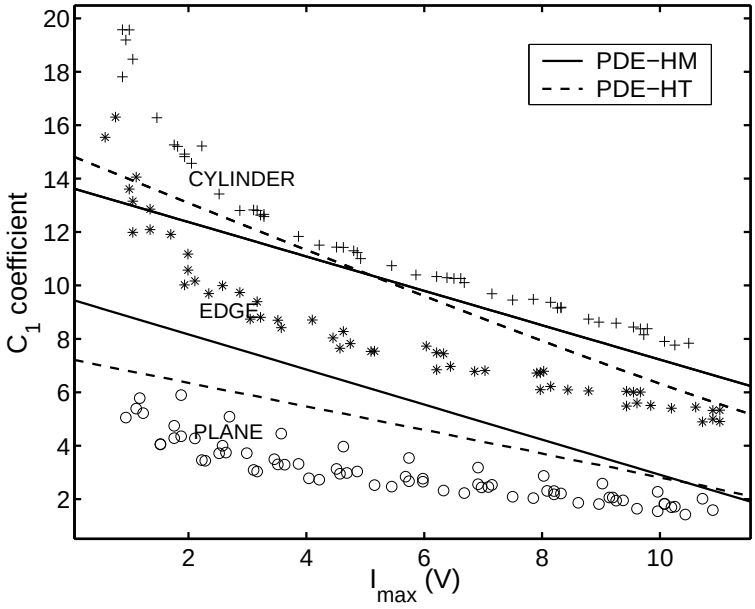



Figure 7.7: Discriminant functions for PDE when the $[C_1, I_{\max}]^T$ feature vector is used.

only three edges are incorrectly classified as cylinders. For heteroscedastic PDE (Table 7.4), the differentiation rate on the training set improves to 98.3% and the correct differentiation rate on the test set is the same as in the homoscedastic case. These results are much better than those obtained with the classification based on the $[C_0, I_{\max}]^T$ feature vector. We have also considered the use of feature vectors $[C_0, C_1, I_{\max}]^T$ and $[C_0, C_1]^T$. However, these did not bring any improvement over those reported. Since the results indicate the $C_1$ parameter is more distinctive then $C_0$ in identifying the geometry, from now on, we concentrate on differentiation based only on the $[C_1, I_{\max}]^T$ feature vector.

Table 7.3: Confusion matrix: homoscedastic PDE using the $[C_1, I_{\max}]^T$ feature vector.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | –(–) | 49(40) | 6(3) | 55(43) |
| CY | –(–) | –(–) | 50(84) | 50(84) |
| total | 70(84) | 49(40) | 56(87) | 175(211) |

Table 7.4: Confusion matrix: heteroscedastic PDE using the $[C_1, I_{\max}]^T$ feature vector.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | –(–) | 52(40) | 3(3) | 55(43) |
| CY | –(–) | –(–) | 50(84) | 50(84) |
| total | 70(84) | 52(40) | 53(87) | 175(211) |

**7.1.1.1.2 Mixture of Normals (MoN) Classifier:** In the MoN classifier, each feature vector in the training set is assumed to be associated with a mixture of $M$ different and independent normal distributions [93]. Each normal distribution has probability density function $p_j$ with mean vector $\boldsymbol{\mu}_j$ and covariance matrix $\boldsymbol{\Sigma}_j$:

$$p_j(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) = \frac{1}{(2\pi)^{(d/2)}|\boldsymbol{\Sigma}_j|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_j)^T\boldsymbol{\Sigma}_j^{-1}(\mathbf{x}-\boldsymbol{\mu}_j)\right], j = 1, \ldots, M \quad (7.2)$$

The $M$ normal distributions are mixed according to the following model, using the mixing coefficients $\alpha_j$:

$$p(\mathbf{x}|\boldsymbol{\Theta}) = \sum_{j=1}^{M} \alpha_j p_j(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) \quad (7.3)$$

Here, $\boldsymbol{\Theta} = [\alpha_1, \ldots, \alpha_M;\ \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_M;\ \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_M]$ is a parameter vector which consists of three sets of parameters and conveniently represents the relevant parameters for the normals to be mixed. The mixing coefficients should satisfy the normalization condition $\sum_{j=1}^{M} \alpha_j = 1$ and $0 \leq \alpha_j \leq 1\ \forall j$ and can be

thought of as prior probabilities of each mixture component so that $\alpha_j$ = Prob{$j'$th component} = $p(j)$ and $\sum_{j=1}^{M} p(j|\mathbf{x}, \mathbf{\Theta}) = 1$. In our implementation, $M$ takes the values two and three. For the $i$'th class, the parameter vector $\mathbf{\Theta}_i$ maximizing Equation (7.3) needs to be estimated, corresponding to the MLE. Since deriving an analytical expression for the MLE is not possible in this case, $\mathbf{\Theta}_i$ is estimated by using expectation-maximization (E-M) clustering which is iterative [88]. The elements of the parameter vector $\mathbf{\Theta}_i$ are updated recursively as follows:

$$\begin{aligned}
\alpha_{ijk} &= \frac{1}{N_i} \sum_{n=1}^{N_i} p(j|\mathbf{x}_n, \mathbf{\Theta}_{i,k-1}) \\
\boldsymbol{\mu}_{ijk} &= \frac{\sum_{n=1}^{N_i} \mathbf{x}_n p(j|\mathbf{x}_n, \mathbf{\Theta}_{i,k-1})}{\sum_{n=1}^{N_i} p(j|\mathbf{x}_n, \mathbf{\Theta}_{i,k-1})} \\
\mathbf{\Sigma}_{ijk} &= \frac{\sum_{n=1}^{N_i} (\mathbf{x}_n - \boldsymbol{\mu}_{ijk})(\mathbf{x}_n - \boldsymbol{\mu}_{ijk})^T p(j|\mathbf{x}_n, \mathbf{\Theta}_{i,k-1})}{\sum_{n=1}^{N_i} p(j|\mathbf{x}_n, \mathbf{\Theta}_{i,k-1})} \quad \text{where} \quad i = 1, \ldots, c \quad \text{and} \quad j = 1, \ldots, M
\end{aligned} \tag{7.4}$$

Here, $\mathbf{\Theta}_{i,k}$ is the parameter vector estimate of the $i$'th class at the $k$'th iteration step and $N_i$ is the number of feature vectors in the training set representing the $i$'th class. The expectation and maximization steps are performed simultaneously. The algorithm proceeds by using the newly derived parameters as the guess for the next iteration. With E-M clustering, even if the dimensionality of the feature vectors increases, fast and reliable parameter estimation can be accomplished.

Each class is considered independent from the others and training is performed separately for each class. For this reason, addition of new classes can be done conveniently by adding the corresponding feature vectors to the training data set and estimating the corresponding class parameter vector.

After estimating the parameter vectors for each class based on the training set feature vectors, testing is done as follows: A target with a given test feature vector $\mathbf{x}$ is assigned to the class whose parameter vector $\mathbf{\Theta}_i$ maximizes Equation (7.3) so that $p(\mathbf{x}|\mathbf{\Theta}_i) > p(\mathbf{x}|\mathbf{\Theta}_l) \quad \forall i \neq l$. Then, the target is labeled as a member of class $w_i$.

The discriminant functions for classification based on $[C_1, I_{\max}]^T$ feature vector are shown in Figure 7.8. Differentiation results for $M = 3$ are given in Table 7.5 in the form of a confusion matrix. For both $M = 2$ and $M = 3$, all training targets are correctly classified using the $[C_1, I_{\max}]^T$ feature vector. In the tests,

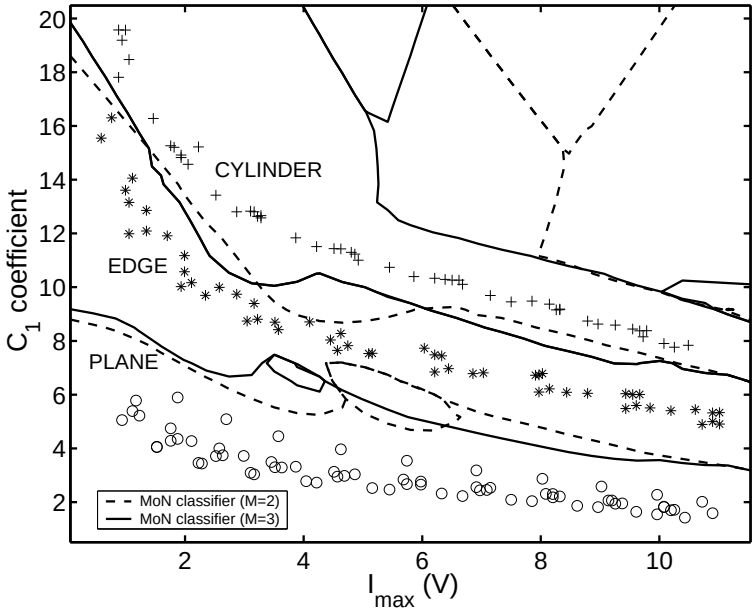



Figure 7.8: Discriminant functions for the MoN classifier when the $[C_1, I_{\max}]^T$ feature vector is used.

for the $M = 3$ case (Table 7.5, in parentheses), again 100% correct differentiation rate is achieved. For the $M = 2$ case, the only difference in the test results is that one of the edges is misclassified as a cylinder so that the correct classification rate falls to 99.5%.

#### 7.1.1.2 Linear and Quadratic Classifiers

**7.1.1.2.1 Linear Classifier by Karhunen Lóeve (KL) Expansion of Covariance Matrix:** This classifier is based on the KL expansion of the common covariance matrix of $c$ classes. The mean vector $\boldsymbol{\mu}$ and the common covariance matrix $\boldsymbol{\Sigma}$ are computed for the training set. The eigenvectors and eigenvalues of the common covariance matrix are computed using the following equations:

$$\begin{aligned} &\boldsymbol{\Sigma}\mathbf{e}_i = \lambda_i \mathbf{e}_i, \qquad i = 1, \ldots, d \\ &(\boldsymbol{\Sigma} - \lambda_i \mathbf{I})\mathbf{e}_i = \mathbf{0} \end{aligned} \tag{7.5}$$

Table 7.5: Confusion matrix: MoN classifier ($M = 3$) using the $[C_1, I_{\max}]^T$ feature vector.

| geometry | diff. result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | –(–) | 55(43) | –(–) | 55(43) |
| CY | –(–) | –(–) | 50(84) | 50(84) |
| total | 70(84) | 55(43) | 50(84) | 175(211) |

Table 7.6: Confusion matrix: linear classifier by KL expansion of the common covariance matrix.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | –(–) | 49(40) | 6(3) | 55(43) |
| CY | –(–) | –(–) | 50(84) | 50(84) |
| total | 70(84) | 49(40) | 56(87) | 175(211) |

where $\mathbf{I}$ is the $2 \times 2$ identity matrix and $\mathbf{0}$ is the $2 \times 1$ zero vector. Eigenvectors with the largest eigenvalues are selected [77]. Although the size of the common covariance matrix ($2 \times 2$) is small for our case, we investigate the effect by either taking the eigenvector with the larger eigenvalue or both of the eigenvectors. Then, the training set is projected onto a subspace of size determined by the number of eigenvectors selected. This is done by $\mathbf{A}^T(\mathbf{x} - \boldsymbol{\mu})$, where the columns of matrix $\mathbf{A}$ consists of the selected eigenvectors. For the one eigenvector case, the correct differentiation rates for training and test sets are 49.7% and 64%, respectively. As expected, the differentiation rates are low for this case. Detailed results are given for the two eigenvector case in Table 7.6. An average correct differentiation rate of 96.7% is achieved on the training set. For test targets, 98.6% is the correct differentiation rate, which is better than that obtained for the training targets.

Table 7.7: Confusion matrix: logistic linear classifier.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | –(–) | 52(40) | 3(3) | 55(43) |
| CY | –(–) | –(–) | 50(84) | 50(84) |
| total | 70(84) | 52(40) | 53(87) | 175(211) |

**7.1.1.2.2 Logistic Linear Classifier:** In the logistic linear classifier, the linear classifier is computed by maximizing the likelihood criterion using the logistic (sigmoid) function [94]. For the two-class problem, logistic classifier maximizes [12]:

$$max_{\theta}\{\prod_{\mathbf{x}_i \in \mathbf{W}_1} \mathbf{q}_1(\mathbf{x}_i^1;\theta) \prod_{\mathbf{x}_i \in \mathbf{W}_2} \mathbf{q}_2(\mathbf{x}_i^2;\theta)\} \tag{7.6}$$

For any discriminant function, $f(\mathbf{x};\theta)$, logistic functions are

$$\begin{aligned} q_1(\mathbf{x};\theta) &= (1+e^{-f(\mathbf{x};\theta)})^{-1}, \\ q_2(\mathbf{x};\theta) &= (1+e^{f(\mathbf{x};\theta)})^{-1} \end{aligned} \tag{7.7}$$

For linear discriminant functions $f(\mathbf{x};\theta)$, Equation (7.6) can be easily optimized. Detailed results are given in Table 7.7. An average correct differentiation rate of 98.3% is achieved on the training set. For test targets, 98.6% is the correct differentiation rate, which is better than that obtained for the training targets.

**7.1.1.2.3 Fisher's Linear Classifier:** Fisher's least-squares linear classifier [77, 94] finds the linear discriminant function between the classes by minimizing the errors in the least-squares sense. The aim is to project data from $d$ dimensions onto a line and find the orientation of the line, [$\mathbf{w}$ in Equation (7.8)], such that the projected data are well separated [77]. The projection is achieved using the following equation:

$$y = \mathbf{w}^T\mathbf{x}, \tag{7.8}$$

Table 7.8: Confusion matrix: Fisher's least-squares linear classifier.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | 18(25) | –(–) | 37(18) | 55(43) |
| CY | –(–) | –(–) | 50(84) | 50(84) |
| total | 88(109) | –(–) | 87(102) | 175(211) |

where $\mathbf{x}$ is the data consisting of $N$ $d$-dimensional samples and $\|\mathbf{w}\| = 1$.

The separation of the projected data points are calculated by finding class means, $\boldsymbol{\mu}_i$. The mean of the projected data is $\tilde{\boldsymbol{\mu}}_i = \mathbf{w}^T\boldsymbol{\mu}_i$. The separation distance between means is $|\tilde{\boldsymbol{\mu}}_1 - \tilde{\boldsymbol{\mu}}_2| = |\mathbf{w}^T(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)|$. Fisher linear classifier uses a linear function $\mathbf{w}^T\mathbf{x}$ such that the criterion given by the following equation is maximized:

$$J(\mathbf{w}) = \frac{|\tilde{\boldsymbol{\mu}}_1 - \tilde{\boldsymbol{\mu}}_2|^2}{|\tilde{\boldsymbol{s}}_1 - \tilde{\boldsymbol{s}}_2|^2}, \tag{7.9}$$

where $\tilde{\boldsymbol{s}}_i$ is scatter of class $w_i$ and computed as $\tilde{\boldsymbol{s}}_i^2$. For a $c$ class problem, there will be $c-1$ discriminant functions, where the dimension of the data should be greater than or equal to the number of classes. Details for the generalization to multiple classes can be found in [77]. The correct differentiation rates are lower than the previous cases, where 68.6% and 79.6% correct differentiation rates are obtained for the training and test targets, respectively. This is due to overlapping of the training data. Detailed results are given in Table 7.8.

**7.1.1.2.4 Nearest Mean Classifier:** We also applied nearest mean classifier and its scaled version, where the linear discriminant function is computed assuming zero covariances and equal class variances. The differentiation rates are low as expected from the distribution of the classes since the data from different classes are correlated. For the nearest mean classifier case, the correct differentiation rates are 82% and 75% for the training and test targets, respectively. For

Table 7.9: Confusion matrix: quadratic discriminant classifier.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | –(–) | 51(40) | 4(3) | 55(43) |
| CY | –(–) | –(–) | 50(84) | 50(84) |
| total | 70(84) | 51(40) | 54(87) | 175(211) |

the scaled version, the respective rates are 82.2% and 75.4%.

**7.1.1.2.5 Quadratic Discriminant Classifier:** In this case, covariance matrix is computed according to the following equation:

$$\mathbf{\Sigma} = (1-R-S) \cdot \mathbf{\Sigma} + R \cdot diag(diag(\mathbf{\Sigma})) + S \cdot mean(diag(\mathbf{\Sigma})) \cdot eye(size(\mathbf{\Sigma}, 1)) \quad (7.10)$$

Regularization parameters ($R$ and $S$) can take values between 0 and 1. *diag* function returns diagonal values of a given matrix or constitutes a matrix, whose diagonal values are the given vector, *eye* returns identity matrix, and *size* returns size of the given matrix. In the first case, regularization parameters are set to zero. The correct differentiation rates are 98.6% and 97.7% for the training and test targets, respectively. Detailed results are given in Table 7.9. When $R$ and $S$ are set to 0.5 for example, the respective rates become 74.9% and 81.5%, which is lower than the previous case.

### 7.1.1.3 Non-Parametric Classifiers

In this section, we consider different non-parametric classifiers, which are the kernel estimator, $k$-nearest neighbor, artificial neural network, and support vector machine classifiers.

**7.1.1.3.1 Kernel Estimator (KE):** KE is a family of PDF estimators first proposed by Fix and Hodges in 1951 [95]. In the KE method, the CCPDF estimates $\hat{p}(\mathbf{x}|w_i)$ are of the form

$$\hat{p}(\mathbf{x}|w_i) = \frac{1}{N_i h_i^d} \sum_{n=1}^{N_i} K\left(\frac{\mathbf{x} - \mathbf{x}_n}{h_i}\right) \quad i = 1, \ldots, c \tag{7.11}$$

where $\mathbf{x}$ is the $d$-dimensional feature vector at which the estimate is being made and $\mathbf{x}_n,\ n = 1, \ldots, N_i$ are the training set sample feature vectors associated with class $w_i$. Here, $h_i$ is called the *spread* or *smoothing parameter* or the *bandwidth* of the KE, and $K(\mathbf{z})$ is a kernel function which satisfies the conditions $K(\mathbf{z}) \geq 0$ and $\int K(\mathbf{z}) d\mathbf{z} = 1$. In this method, the selection of the bandwidth $h_i$ is important [96, 97]. If $h_i$ is selected too small, $\hat{p}(\mathbf{x}|w_i)$ degenerates into a collection of $N_i$ sharp peaks, each located at a sample feature vector. On the other hand, if $h_i$ is selected too large, the estimate is oversmoothed and an almost uniform CCPDF results. Usually, $h_i$ is chosen as a function of $N_i$ such that $\lim_{N_i \to \infty} h(N_i) = 0$.

In the implementation of this method, since $d = 2$, we employed a 2-dimensional normal kernel function. The bandwidth $h_i$ for the $i$th class is precomputed based on the $N_i$ sample feature vectors available for all classes by optimization with respect to *leave-one-out* error [88]. After $h_i$'s are computed, a test feature vector $\mathbf{x}$ is classified into that class for which the CCPDF in Equation (7.11) is maximized. This requires the training data to be stored throughout testing.

**7.1.1.3.2 $k$-Nearest Neighbor ($k$-NN) Classifier:** Consider the $k$ nearest neighbors of a feature vector $\mathbf{x}$ in a set of several feature vectors. Suppose $k_i$ of these $k$ vectors come from class $w_i$. Then, a $k$-NN estimator for class $w_i$ can be defined as $\hat{p}(w_i|\mathbf{x}) = \frac{k_i}{k}$, and $\hat{p}(\mathbf{x}|w_i)$ can be obtained from $\hat{p}(\mathbf{x}|w_i)\hat{p}(w_i) = \hat{p}(w_i|\mathbf{x})\hat{p}(\mathbf{x})$. This results in a classification rule such that $\mathbf{x}$ is classified into class $w_j$ if $k_j = \max_i(k_i)$, where $i = 1, \ldots, c$. In other words, the $k$ nearest neighbors of the vector $\mathbf{x}$ in the training set are considered and the vector $\mathbf{x}$ is classified into the same class as the majority of its $k$ nearest neighbors.

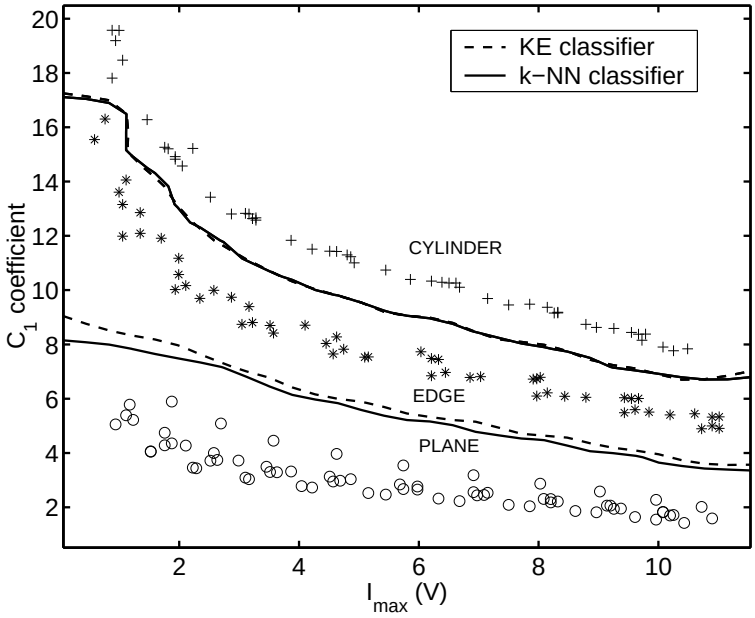



Figure 7.9: Discriminant functions for the KE and the $k$-NN classifier when the $[C_1, I_{\max}]^T$ feature vector is used.

A major disadvantage of this method is that a pre-defined rule for the selection of the value of $k$ does not exist [98]. In this study, the number of nearest neighbors $k$ is determined by optimization with respect to leave-one-out error. In the implementation of the $k$-NN classifier, $k$ values varying between 1 and 12 have been considered. For $k = 1, 2$ and 3, the same correct differentiation rates are obtained (given below) for the training and test sets, respectively. For larger values of $k$, the errors start increasing. Again, the training data must be stored during testing.

In Figure 7.9, the discriminant functions for both the KE and the $k$-NN classifiers are illustrated when the $[C_1, I_{\max}]^T$ feature vector is used. For both classifiers, the training targets are correctly differentiated with 100% correct differentiation rate. For the test targets, only one edge target is incorrectly classified as a cylinder, corresponding to a correct differentiation rate of 99.5%.

Table 7.10: Confusion matrix: ANN trained with BP.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | –(–) | 52(40) | 3(3) | 55(43) |
| CY | –(–) | –(–) | 50(84) | 50(84) |
| total | 70(84) | 52(40) | 53(87) | 175(211) |

**7.1.1.3.3 Artificial Neural Network (ANN) Classifiers:** Feed-forward ANNs trained with back-propagation (BP) and Levenberg-Marquardt (LM) algorithms, and a linear perceptron (LP) are used as classifiers. The feed-forward ANN has one hidden layer with four neurons. The number of neurons in the input layer is two (since the feature vector consists of two parameters) and the number of neurons in the output layer is three. LP is the simplest type of ANN, used for classification of two classes that are linearly separable. LP consists of a single neuron with adjustable input weights and a threshold value [99]. If the number of classes is greater than two, LPs are used in parallel. One perceptron is used for each output. The maximum number of epochs is chosen as 1000. The weights are initialized randomly and the learning rate is chosen as 0.1. MATLAB$^{TM}$Neural Network Toolbox is used for the implementation. The discriminant functions are given in Figure 7.10 for the three classifiers. The correct differentiation rates using the BP algorithm are given in Table 7.10. Differentiation rates of 98.3% and 98.6% are achieved for the training and test sets, respectively. When training is done by LM, the same correct differentiation rate is obtained on the training set (see Table 7.11). However, this classifier is better than the BP method in the tests, where only one edge target is misclassified as a cylinder, resulting in a correct differentiation rate of 99.5%. The results for the LP classifier are given in Table 7.12. As expected from the distribution of the parameters, because the classes are not linearly separable, lower correct differentiation rates of 77.7% and 76.3% are achieved on the training and test sets, respectively.

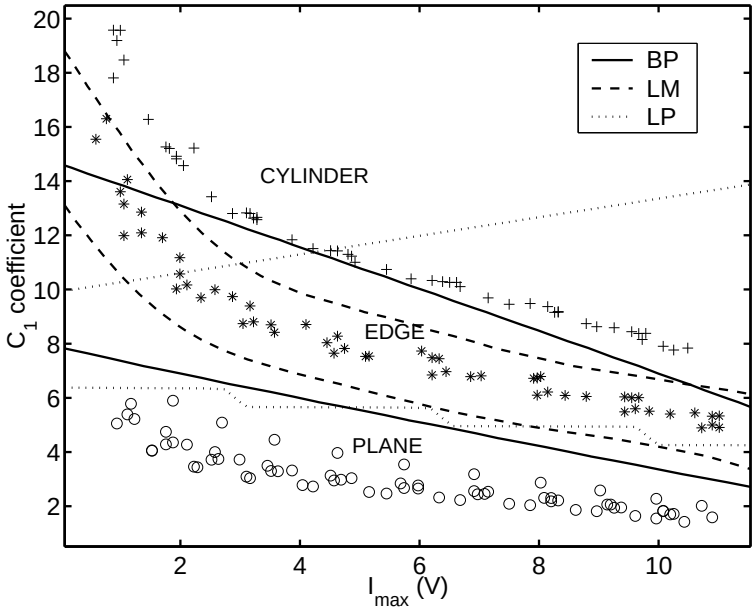



Figure 7.10: Discriminant functions for ANN classifiers when the $[C_1, I_{\max}]^T$ feature vector is used.

**7.1.1.3.4 Support Vector Machine (SVM) Classifier:** SVM classifier is a machine learning technique proposed early in the eighties [100]. It has been used in applications such as object, voice, and handwritten character recognition, and text classification.

If the feature vectors in the original feature space are not linearly separable, SVMs preprocess and represent them in a space of higher dimension where they become linearly separable. The dimension of the transformed space is typically

Table 7.11: Confusion matrix: ANN trained with LM.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | –(–) | 52(42) | 3(1) | 55(43) |
| CY | –(–) | –(–) | 50(84) | 50(84) |
| total | 70(84) | 52(42) | 53(85) | 175(211) |

Table 7.12: Confusion matrix: ANN trained with LP.

| geometry | differentiation result | | | total |
|---|---|---|---|---|
| | P | E | CY | |
| P | 70(84) | –(–) | –(–) | 70(84) |
| E | –(–) | 45(32) | 10(11) | 55(43) |
| CY | –(–) | 29(39) | 21(45) | 50(84) |
| total | 70(84) | 74(71) | 31(56) | 175(211) |

much higher than the original feature space. With a suitable nonlinear mapping $\phi(.)$ to a sufficiently high dimension, data from two different classes can always be made linearly separable, and separated by a hyperplane [101]. The choice of the nonlinear mapping depends on the prior information available to the designer. If such information is not available, one might choose to use polynomials, Normals, or other types of basis functions. The dimensionality of the mapped space can be arbitrarily high. However, in practice, it may be limited by computational resources. The complexity of SVMs is related to the number of resulting support vectors rather than the high dimensionality of the transformed space.

Consider SVMs in a binary classification setting. We are given the training feature vectors $\mathbf{x}_i$ that are vectors in some space $\mathcal{X} \subseteq \Re^d$ and their labels $l_i \in \{-1, 1\}$ where $i = 1, \ldots, N$. The goal in training a SVM is to find the separating hyperplane with the largest margin so that the generalization of the classifier is better. All vectors lying on one side of the hyperplane are labeled as +1, and all vectors lying on the other side are labeled as –1. The support vectors are the (transformed) training patterns that lie closest to the hyperplane and are at equal distance from it. They correspond to the training samples that define the optimal separating hyperplane and are the most difficult patterns to classify, yet the most informative for the classification task.

More generally, SVMs allow one to project the original training data in space $\mathcal{X}$ to a higher dimensional feature space $\mathcal{F}$ via a Mercer kernel operator $K$ [102]. We consider a set of classifiers of the form $f(\mathbf{x}) = \sum_{i=1}^{N} \beta_i \; K(\mathbf{x}_i, \mathbf{x})$. When $f(\mathbf{x}) \geq 0$, we label $\mathbf{x}$ as +1, otherwise as –1. When $K$ satisfies Mercer's condition,

$K(\mathbf{u}, \mathbf{v}) = \phi(\mathbf{u}) \cdot \phi(\mathbf{v})$ where $\phi(.) : \mathcal{X} \rightarrow \mathcal{F}$ is a nonlinear mapping and "$\cdot$" denotes the inner product. We can then rewrite $f(\mathbf{x})$ as $f(\mathbf{x}) = \mathbf{a} \cdot \phi(\mathbf{x})$, where $\mathbf{a} = \sum_{i=1}^{N} \beta_i \ \phi(\mathbf{x}_i)$ is a weight vector. Thus, by using $K$, the training data is projected into a new feature space $\mathcal{F}$ which is often higher dimensional. The SVM then computes the $\beta_i$'s that correspond to the maximal margin hyperplane in $\mathcal{F}$. By choosing different kernel functions, we can project the training data from $\mathcal{X}$ into spaces $\mathcal{F}$ for which hyperplanes in $\mathcal{F}$ correspond to more complex decision boundaries in the original space $\mathcal{X}$. Hence, by nonlinear mapping of the original training patterns into other spaces, decision functions can be found using a linear algorithm in the transformed space by only computing the kernel $K(\mathbf{x}_i, \mathbf{x})$.

The function $f(\mathbf{x}) = \mathbf{a} \cdot \phi(\mathbf{x})$ is a linear discriminant function in the transformed space based on the hyperplane $\mathbf{a} \cdot \phi(\mathbf{x}) = 0$. Here, both the weight vector and the transformed feature vector have been augmented by one dimension to include a bias weight so that the hyperplanes need not pass through the origin. A separating hyperplane ensures

$$l_i \ f(\mathbf{x}_i) = l_i \ \mathbf{a} \cdot \phi(\mathbf{x}_i) \geq 1 \quad \text{for} \quad i = 1, \ldots, N \tag{7.12}$$

It can be shown that finding the optimal hyperplane corresponds to minimizing the magnitude of the weight vector $\| \ \mathbf{a} \ \|^2$ subject to the constraint given by Equation (7.12) [77]. Using the method of Lagrange multipliers, we construct the functional

$$L(\mathbf{a}, \lambda) = \frac{1}{2} \| \ \mathbf{a} \ \|^2 - \sum_{i=1}^{N} \lambda_i \ [l_i \ \mathbf{a} \cdot \phi(\mathbf{x}_i) - 1] \tag{7.13}$$

where the second term in the above equation expresses the goal of classifying the points correctly. To find the optimal hyperplane, we minimize $L(.)$ with respect to the weight vector $\mathbf{a}$, while maximizing with respect to the undetermined Lagrange multipliers $\lambda_i \geq 0$. This can be done by solving the constrained optimization problem by quadratic programming [88] or by other alternative techniques. The solution of the weight vector is $\mathbf{a}^* = \sum_{i=1}^{N} l_i \ \lambda_i \ \phi(\mathbf{x}_i)$ corresponding to $\beta_i = l_i \lambda_i$. Then the decision function is given by:

$$f^*(\mathbf{x}) = \sum_{i=1}^{N} \lambda_i \ l_i \ \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}) \tag{7.14}$$

In this study, the method described above is applied to differentiate target feature vectors from multiple classes. Following the one-versus-rest method, $c$ different binary classifiers are trained, where each classifier recognizes one of $c$ target types. SVM classifiers with polynomial, exponential, and radial basis function kernels are used. The kernel functions are $K_p(\mathbf{x}, \mathbf{x}_i) = (\mathbf{x} \cdot \mathbf{x}_i + 1)^3$, $K_e(\mathbf{x}, \mathbf{x}_i) = e^{-\|\mathbf{x}-\mathbf{x}_i\|}$, $K_r(\mathbf{x}, \mathbf{x}_i) = e^{-\|\mathbf{x}-\mathbf{x}_i\|^2}$, respectively [88]. The dimension of the feature space $\mathcal{F}$ is 3. 100% correct differentiation rate is achieved on the training set for all of the SVM classifiers. For the test set, the correct differentiation rates are 99.5%, 99.5%, and 99.1% for SVM classifiers with polynomial, exponential, and radial basis function kernels, respectively. Therefore, the polynomial and exponential kernels result in the highest classification rates.

To summarize the results of the statistical pattern recognition techniques for geometry classification based on the $[C_1, I_{\max}]^T$ feature vector, the overall differentiation rates are given in Table 7.13. Best classification rate is obtained for the test scans using the MoN classifier with three components. This is followed by MoN with two components, KE, $k$-NN, and SVM with polynomial and exponential kernels, equally. Ranking according to highest classification rate continues as ANN trained with LM algorithm, SVM with radial kernel, heteroscedastic and homoscedastic PDE, KL, logistic classifier, ANN trained with BP, SVM with polynomial kernel, quadratic discriminant classifier, Fisher's linear classifier, ANN trained with LP, scaled nearest mean classifier, and nearest mean classifier.

### 7.1.2 Determination of Surface Type

Parametric surface differentiation is a more difficult problem than geometry differentiation. This is clearly seen in the very similar variation of the parameters for different surfaces corresponding to the same geometry (Figure 7.4). In [60], planar surfaces covered with six different surfaces are correctly classified with 100% correct differentiation rate. Although we succeeded with surface differentiation for planar surfaces, the surface differentiation results for other geometries were not as good. The above classification approaches were applied to differentiate between surface types assuming the geometry of the targets is determined

Table 7.13: Correct differentiation percentages for different classifiers (PDE-HM: Parametric density estimation-homoscedastic, PDE-HT: Parametric density estimation-heteroscedastic, LC-KL: Linear classifier-Karhunen Lóeve, LC-LOG: Linear classifier-logistic, LC-FIS: Linear classifier-Fisher's least-squares, NM: nearest mean classifier, NMS: nearest mean scaled classifier, QC: quadratic classifier, MoN-2: Mixture of normals with two components, MoN-3: Mixture of normals with three components, KE: kernel estimator, $k$-NN: $k$-nearest neighbor, ANN-BP: ANN trained with BP, ANN-LM: ANN trained with LM, ANN-LP: ANN trained with LP, SVM-P: SVM with polynomial kernel, SVM-E: SVM with exponential kernel, SVM-R: SVM with radial kernel).

| classification techniques | data set | |
|---|---|---|
| | training | test |
| PDE-HM | 96.6 | 98.6 |
| PDE-HT | 98.3 | 98.6 |
| LC-KL | 96.7 | 98.6 |
| LC-LOG | 98.3 | 98.6 |
| LC-FIS | 68.6 | 79.6 |
| NM | 82 | 75 |
| NMS | 82.2 | 75.4 |
| QC | 98.6 | 97.7 |
| MoN-2 | 100 | 99.5 |
| MoN-3 | 100 | 100 |
| KE | 100 | 99.5 |
| $k$-NN | 100 | 99.5 |
| ANN-BP | 98.3 | 98.6 |
| ANN-LM | 98.3 | 99.5 |
| ANN-LP | 77.7 | 76.3 |
| SVM-P | 100 | 99.5 |
| SVM-E | 100 | 99.5 |
| SVM-R | 100 | 99.1 |

correctly beforehand. For example, for cylindrical targets, the classification error is about 85% when PDE is used. Since the results were not promising, no further attempt has been made to differentiate surface types in parametric space.

As an alternative, we extracted features from these intensity scans corresponding to different surfaces of the same geometry using *forward feature selection*. Since the magnitude and basewidth of intensity scans both change with distance, the intensity scans are first normalized before feature extraction. We experimented with different features of the intensity scans by extracting the points representing the intensity scans best and using them for differentiation. However, the differentiation results were not promising. For example, for cylindrical targets, the surfaces are correctly classified only with a correct differentiation rate of 20%. Different initialization procedures did not result in any improvement in feature extraction.

In this chapter, we extended the parametric surface differentiation approach proposed in [60] to differentiate both the geometry and surface type of the targets using statistical pattern recognition techniques. We compared different classifiers such as PDE, LC-KL, LC-LOG, LC-FIS, NM, NMS, QC, MoN, kernel estimator, $k$-NN, ANN, and SVM for geometry type determination. Best differentiation rates (100%) are obtained for the MoN classifier with three components. MoN classifier performs better than models which associate the data with a single distribution. It is also more robust and the training set can be easily updated when new classes need to be added to the database.

In the next chapter, a comparison of the proposed methods is made.

# Chapter 8

# COMPARISON OF THE TECHNIQUES

This chapter provides a summary of the performances of the different differentiation approaches used throughout this thesis for target differentiation and localization of commonly encountered features in indoor environments. To the best of our knowledge, no attempt has been made to differentiate and estimate the position of several kinds of targets using simple infrared sensors. Also, a compact comparison based on experimental data does not exist for target differentiation using infrared sensors. Differentiation results of each approach were given at the end of the corresponding chapter. Here, we will summarize the results and provide a compact comparison.

Rule-based approach, described briefly in Chapter 3, achieves position-invariant target differentiation without relying on the absolute return signal intensities of the infrared sensors. The target primitives employed in the rule-based approach are plane, corner, edge, and cylinder, all made of unpolished oak wood. An average correct differentiation rate of 91.3% is achieved.

In the template-based approach, discussed in Chapter 4, an average correct classification rate of 93% is obtained with the least-squares approach for targets with different geometrical properties (plane, corner, edge, and cylinder) but made

of the same surface material (wood). For the matched filtering case, the average correct differentiation rate over all target types is 97%, which is better than that obtained with the least-squares approach. For different surface materials (aluminum, white wall, brown paper, and Styrofoam) of the same planar geometry, the average correct classification rate obtained by using the least-squares approach is 82%. For the matched filtering case, the average correct differentiation rate over all surfaces is 87%. For targets with both different geometry and surface properties (plane, corner, and edge covered with aluminum, white cloth, and Styrofoam), 77% average correct classification rate is achieved by using least-squares. With matched filter, the average accuracy of differentiation over all target types is 80%.

Results using artificial neural networks are given in Chapter 5. The training algorithms employed are BP and LM. The network trained with LM and pruned with Optimal Brain Surgeon technique gives differentiation results which are comparable with those obtained with template-based target differentiation, where geometry type of the targets is classified with 99% accuracy and an overall correct differentiation rate of 78.4% is achieved for all surfaces.

The parametric surface differentiation approach is discussed in Chapter 6. For a set of six surfaces including Styrofoam packaging material, white painted matte wall, white or black cloth, and white, brown, and violet paper (also matte), we get a correct differentiation rate of 100%.

For the statistical pattern recognition techniques (Chapter 7), mixture of normals classifier with three components correctly differentiates three types of geometries with different surface properties, resulting in the best performance (100%) in geometry differentiation.

Table 8.1 summarizes the results for all of the methods considered, allowing their overall comparison. In this summary table, the first column represents the methods used in the classification. References to our related publications are given in this column for more detail. The second and third columns represent the geometries and surfaces considered for each classification method. The result type indicates the differentiation of geometry and/or surface. Best differentiation

rates are given for the different variations of the methods considered.

The matched filtering approach gives better results in the template-based differentiation. We have seen that the geometrical properties of the targets are more distinctive than their surface properties, and surface determination is the limiting factor. Based on the data we have collected, it seems possible to increase the vocabulary of different geometries, provided they are not too similar. However, the same cannot be said for the number of different surfaces. For a given total number of distinct targets, increasing the number of surfaces and decreasing the number of geometries will in general make the results worse. On the other hand, decreasing the number of surfaces and increasing the number of geometries will in general improve the results. The method we propose as a template-based differentiation is scalable in the sense that the accuracy can be increased by increasing the number of reference scans without increasing the computational cost.

The differentiation results obtained using artificial neural networks are comparable with those obtained in template-based differentiation. Planes and corners covered with aluminum are correctly classified in all of our studies due to their distinctive features. In both approaches, the greatest difficulty is encountered in the differentiation of edges of different surface types.

The parametric approach can differentiate six different surfaces with 100% accuracy. In the template-based approach, where we considered differentiation and localization of surfaces by employing non-parametric approaches, a maximum correct differentiation rate of 87% over four surfaces was achieved. Comparing this rate with that obtained with the parametric approach, we can conclude that the parametric approach is superior to non-parametric ones, in terms of the accuracy, number of surfaces differentiated, and memory requirements, since the non-parametric approaches we considered require the storage of reference scan signals. By parameterizing the intensity scans and storing only their parameters, we have eliminated the need to store complete reference scans. The decrease in the differentiation rate resulting from adding new surfaces in the parametric approach does not represent an overall degradation in differentiation rates across all surface types but is almost totally explained by pairwise confusion of the newly

Table 8.1: Overview of the differentiation techniques compared (U: used, S: stored, and NS: not stored).

| differentiation technique | type of geometry | type of surface | feature | correct diff.(%) | training data | learning | parametric |
|---|---|---|---|---|---|---|---|
| rule-based [56] | P,C,E,CY | WD | geo | 91.3 | U, NS | no | no |
| template-based | | | | | | | |
| [57] | P,C,E,CY | WD | geo | 97 | U, S | no | no |
| [58] | P | AL,WW,BP,ST | surf | 87 | ” | ” | ” |
| [59] | P,C,E | AL,WC,ST | geo | 99 | ” | ” | ” |
| ” | ” | ” | surf | 81 | ” | ” | ” |
| ” | ” | ” | geo+surf | 80 | ” | ” | ” |
| ANN | P,C,E | AL,WC,ST | geo | 99.0 | U, NS | yes | no |
| | P | ” | surf | 80.5 | ” | ” | ” |
| | C | ” | ” | 85.9 | ” | ” | ” |
| | E | ” | ” | 64.6 | ” | ” | ” |
| | P,C,E | ” | geo+surf | 78.4 | ” | ” | ” |
| parametric [60] | P | ST,WW,WC(BC), WP,BP,VP | surf | 100 | U, NS | yes | yes |
| | ” | ST,WW,WC(BC), WP,BP,VP,WD | ” | 86 | ” | ” | ” |
| | ” | ST,WW,WC,BC, WP,BP,VP | ” | 83 | ” | ” | ” |
| | ” | ST,WW,WC,BC, WP,BP,VP,WD | ” | 73 | ” | ” | ” |
| statistical pattern recognition [61] | | | | | | | |
| PDE-HM, PDE-HT | P,E,CY | ST,WC,BC, WP,BP,VP,WD | geo | 98.6 | U, NS | no | yes |
| LC-KL | ” | ” | ” | 98.6 | ” | ” | no |
| LC-LOG | ” | ” | ” | 98.6 | ” | ” | ” |
| LC-FIS | ” | ” | ” | 79.6 | ” | ” | ” |
| NM | ” | ” | ” | 75 | ” | ” | ” |
| NMS | ” | ” | ” | 75.4 | ” | ” | ” |
| QC | ” | ” | ” | 97.7 | ” | ” | ” |
| MoN-3 | ” | ” | ” | 100 | ” | ” | yes |
| KE | ” | ” | ” | 99.5 | U, S | ” | no |
| $k$-NN | ” | ” | ” | 99.5 | ” | ” | ” |
| NN-LM | ” | ” | ” | 99.5 | U, NS | yes | ” |
| SVM-P, SVM-E | ” | ” | ” | 99.5 | ” | no | ” |

introduced surface with a previously existing one, resulting from the similarity of the $C_1$ parameter of the intensity scans of the two confused surfaces. (Similar decreases in differentiation rate with increasing number of surfaces or objects are also observed with non-parametric template-based approaches.) We can increase the number of surfaces differentiated at the expense of a decrease in the correct differentiation rate.

Surface differentiation using statistical pattern recognition techniques was not as successful as geometry type determination due to the similar ($I_{max}$,$C_1$) variation of edges and cylinders.

We give localization results for rule-, template-, and parameter-based differentiation. As the approaches for target localization are the same for other classification approaches (ANNs and statistical pattern recognition techniques), results are not given for these cases. Emphasis is made on target differentiation. Two alternatives, center-of-gravity and maximum intensity of the intensity scans, are used for azimuth estimation of the targets depending whether the intensity scans are saturated or not. After determination of the target type, range of the targets is found by interpolating on the intensity versus distance curve. Therefore, the greatest contribution to the range errors comes from targets which are incorrectly differentiated and/or whose intensity scans are saturated.

The correct differentiation rate is low for targets located at far or nearby distances to the infrared sensing unit, as the intensity scan is weak or saturated for the two extreme cases, respectively. The experiments are conducted in a controlled environment, but the data for training and test scans are collected at different times. The results are consistent over time and for different environmental conditions. While performing differentiation and localization, scans for the training sets should be obtained carefully. In our case, targets were located on millimetric grid paper during data acquisition. The orientation of the targets is also done carefully. We also considered cases where the scans deviate from ideal cases. The proposed algorithms can be easily modified with minor modifications for different environments and target/surface types. The methods can be applied/implemented successfully in real-time for on-board applications.

In real-time, a circular array of emitter/detector pairs can be used for fast and on-line target determination. This way, data acquisition can be done faster. The complexity of the data processing for target type determination is low, where the processing of the training set took more time. As training is done off-line for some of the methods, it does not degrade the real-time performance of the proposed algorithm. Rule-based approach processes full intensity scans, and the decision is made only using simple computations. Template-based approach also uses full intensity scans, and processing taking differences with reference scans or matched filtering requires more computational cost than rule-based approach but again it is fast enough for real-time applications. Also, neural network based target differentiation is comparable to the template-based approach in terms of processing time. Other approaches do not use full intensity scans, but only two parameters obtained by fitting a reflection model to the scans. The fitting process is also suitable for real-time applications where model-based clever initial guesses are made for fast convergence.

In this chapter, a comparison of all approaches used for target classification and localization throughout this thesis is made. Because of the different properties of each target, the number of scans per geometry and surface and the range interval where the targets are visible by the experimental setup are not the same. This is in the nature of the application, therefore no attempt has been made to make the number of training scans equal as this will also introduce a bias which inherently is not part of the nature of the application.

The results provided in this thesis are vital for robotics researchers who are looking for which method results in better target classification and localization performance with infrared sensors. In the next chapter, conclusions are drawn and directions for future work are provided.

# Chapter 9

# CONCLUSIONS AND FUTURE WORK

In this thesis, differentiation and localization of commonly encountered indoor features or targets such as planes, corners, edges, and cylinders with different surfaces is achieved using an inexpensive infrared emitter and detector pair. One advantage of our system is that it does not greatly depend on environmental conditions since we employ an active sensing modality.

Different approaches are compared in terms of correct target differentiation rate, and range and azimuth estimation accuracy. The techniques considered in this study include rule-based, template-based and neural network-based differentiation, parametric surface differentiation, and statistical pattern recognition techniques such as parametric density estimation, different linear and quadratic classifiers, mixture of normals, kernel estimator, $k$-nearest neighbor, artificial neural network, and support vector machine classifiers.

The results reported here represent the outcome of our efforts to explore the limits of what is achievable in terms of identifying information with only a simple emitter-detector pair. Such simple sensors are usually put to much lower information extracting uses. To the best of our knowledge, no previous attempt has been made to differentiate and estimate the position of several kinds of targets using

such simple infrared sensors. Also, a compact comparison based on experimental work does not exist for target differentiation using infrared sensors.

This thesis demonstrates that simple infrared sensors, when coupled with appropriate processing, can be used to extract substantially more information than such devices are commonly employed for. We expect this flexibility to significantly extend the range of applications in which such low-cost single sensor based systems can be used. Specifically, we expect that it will be possible to go beyond relatively simple tasks such as simple object and proximity detection, counting, distance and depth monitoring, floor sensing, position measurement, obstacle/collision avoidance, and deal with tasks such as differentiation, classification, recognition, clustering, position estimation, map building, perception of the environment and surroundings, autonomous navigation, and target tracking. The approach presented here would be more useful where self-correcting operation is possible due to repeated observations and feedback.

The demonstrated system would find application in intelligent autonomous systems such as mobile robots whose task involves surveying an unknown environment made of different surface types. Industrial applications where different materials/surfaces must be identified and separated may also benefit from this approach.

Given the attractive performance-for-cost of infrared-based systems, we believe that the results of this study will be useful for engineers designing or implementing infrared systems and researchers investigating algorithms and performance evaluation of such systems. While we have concentrated on infrared sensing, the techniques evaluated and compared in this thesis may be useful for other sensing modalities and environments where the objects are characterized by complex signatures and the information from a multiplicity of partial viewpoints must be combined and resolved.

Future work may involve designing a more intelligent system whose operating range is adjustable based on an initial range estimate to the target. This will eliminate saturation and enable the system to accurately differentiate and localize targets over a wider operating range. Another issue to consider is the extension

of the model to include specular reflections from glossy surfaces. Parametric modeling and representation of intensity scans of different geometries (such as corner, edge, and cylinder) can be considered to employ the proposed approach in the simultaneous determination of the geometry and the surface type of the targets. Identifying more generally shaped objects (such as a vase or a bottle) by using several scans from each object is another possible research direction. In a sensor-fusion framework, infrared sensors would be perfectly complementary to ultrasonic systems which are not suitable for close-range detection (less than 40 cm). They can be used together with ultrasonic sensors for target differentiation purposes [1–6]. Finally, evaluating the techniques on a mobile robot moving in indoor environments would put the system into practical use.

# Bibliography


[1] R. Kuc and M. W. Siegel, "Physically-based simulation model for acoustic sensor robot navigation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, pp. 766–778, November 1987.

[2] L. Kleeman and R. Kuc, "Mobile robot sonar for target localization and classification," *International Journal of Robotics Research*, vol. 14, pp. 295–318, August 1995.

[3] J. J. Leonard and H. F. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," *IEEE Transactions on Robotics and Automation*, vol. 7, pp. 376–382, June 1991.

[4] J. Manyika and H. F. Durrant-Whyte, *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*. New York: Ellis Horwood, 1994.

[5] Ö. Bozma and R. Kuc, "Building a sonar map in a specular environment using a single mobile sensor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 1260–1269, December 1991.

[6] M. L. Hong and L. Kleeman, "Analysis of ultrasonic differentiation of three-dimensional corners, edges and planes," in *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 580–584, Nice, France, 12–14 May 1992.

[7] P. J. Phillips, "Matching pursuit filters applied to face identification," *IEEE Transactions on Image Processing*, vol. 7, pp. 1150–1164, August 1998.

[8] H. Kwon, S. Z. Der, and N. M. Nasrabadi, "Adaptive multisensor target detection using feature-based fusion," *Optical Engineering*, vol. 41, pp. 69–80, January 2002.

[9] T. Tsao and Z. Q. Wen, "Image-based target tracking through rapid sensor orientation change," *Optical Engineering*, vol. 41, pp. 697–703, March 2002.

[10] I. Pavlidis, P. Symosek, B. Fritz, M. Bazakos, and N. Papanikolopoulos, "Automatic detection of vehicle occupants: the imaging problem and its solution," *Machine Vision and Applications*, vol. 11, pp. 313–320, April 2000.

[11] P. M. Tag, R. L. Bankert, and L. R. Brody, "An AVHRR multiple cloud-type classification package," *Journal of Applied Meteorology*, vol. 39, pp. 125–134, February 2000.

[12] A. K. Jain, N. K. Ratha, and S. Lakshmanan, "Object detection using Gabor filters," *Pattern Recognition*, vol. 30, pp. 295–309, February 1997.

[13] Z. Zalevsky, D. Mendlovic, E. Rivlin, and S. Rotman, "Contrasted statistical processing algorithm for obtaining improved target detection performances in infrared cluttered environment," *Optical Engineering*, vol. 39, pp. 2609–2617, October 2000.

[14] B. Bhanu, P. Symosek, and S. Das, "Analysis of terrain using multispectral images," *Pattern Recognition*, vol. 30, pp. 197–215, February 1997.

[15] F. T. S. Yu and S. Jutamulia, ed., *Optical Pattern Recognition*. Cambridge: Cambridge University Press, 1998.

[16] F. T. S. Yu and S. Yin, ed., *Selected Papers on Optical Pattern Recognition*, vol. MS 156 of *SPIE Milestone Series*. Bellingham, Washington: SPIE Optical Engineering Press, 1999.

[17] D. Casasent and D. Psaltis, "Scale invariant optical correlation using Mellin transforms," *Optics Communications*, vol. 17, pp. 59–63, April 1976.

[18] M. McDonnell, "Clarification on use of Mellin transform in optical pattern recognition," *Optics Communications*, vol. 25, no. 3, pp. 320–322, 1978.

[19] H. H. Arsenault, Y. N. Hsu, and K. Chalasinska-Macukow, "Rotation-invariant pattern recognition," *Optical Engineering*, vol. 23, pp. 705–709, November/December 1984.

[20] F. T. S. Yu, X. Li, E. Tam, S. Jutamulia, and D. A. Gregory, "Rotation invariant pattern recognition with a programmable joint transform correlator," *Applied Optics*, vol. 28, pp. 4725–4727, 15 November 1989.

[21] G. Gheen, "Design considerations for low-clutter, distortion invariant correlation filters," *Optical Engineering*, vol. 29, pp. 1029–1032, September 1990.

[22] C. Gu, J. Hong, and S. Campbell, "2-D shift invariant volume holographic correlator," *Optics Communications*, vol. 88, pp. 309–314, 1 April 1992.

[23] A. Lohmann, Z. Zalevsky, and D. Mendlovic, "Synthesis of pattern recognition filters for fractional Fourier processing," *Optics Communications*, vol. 128, pp. 199–204, 15 July 1996.

[24] P. Refregier, "Optical pattern recognition—optimal trade-off circular harmonic filters," *Optics Communications*, vol. 86, pp. 113–118, 1 November 1991.

[25] S. Roy, H. H. Arsenault, and Y. Sheng, "Shift-, scale-, rotation and pose-invariant object recognition using centroid wedge sampling and a feature space trajectory classifier," *Journal of Modern Optics*, vol. 50, pp. 285–297, January 2003.

[26] E. Cheung and V. J. Lumelsky, "Proximity sensing in robot manipulator motion planning: system and implementation issues," *IEEE Transactions on Robotics and Automation*, vol. 5, pp. 740–751, December 1989.

[27] A. J. Hand, "Infrared sensor counts insects," *Photonics Spectra*, vol. 32, pp. 30–31, November 1998.

[28] H. C. Wikle, S. Kottilingam, R. H. Zee, and B. A. Chin, "Infrared sensing techniques for penetration depth control of the submerged arc welding process," *Journal of Materials Processing Technology*, vol. 113, pp. 228–233, 15 June 2001.

[29] B. Butkiewicz, "Position control system with fuzzy microprocessor AL220," *Lecture Notes in Computer Science*, vol. 1226, pp. 74–81, 1997.

[30] V. J. Lumelsky and E. Cheung, "Real-time collision avoidance in teleoperated whole-sensitive robot arm manipulators," *IEEE Transactions on Systems Man and Cybernetics*, vol. 23, pp. 194–203, January/February 1993.

[31] H. R. Everett, *Sensors for Mobile Robots, Theory and Application.* 289 Linden St, Wellesley, MA: A K Peters, Ltd., 1995.

[32] G. Beccari, S. Caselli, and F. Zanichelli, "Qualitative spatial representations from task-oriented perception and exploratory behaviors," *Robotics and Autonomous Systems*, vol. 25, pp. 147–157, 30 November 1998.

[33] A. Warszawski, Y. Rosenfeld, and I. Shohet, "Autonomous mapping system for an interior finishing robot," *Journal of Computing in Civil Engineering*, vol. 10, pp. 67–77, January 1996.

[34] E. P. Lopes, E. P. L. Aude, J. T. C. Silveria, H. Serderia, and M. F. Martins, "Application of a blind person strategy for obstacle avoidance with the use of potential fields," in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 3, pp. 2911–2916, Seoul, South Korea, 21–26 May 2001.

[35] A. M. Flynn, "Combining sonar and infrared sensors for mobile robot navigation," *International Journal of Robotics Research*, vol. 7, pp. 5–14, December 1988.

[36] H. M. Barberá, A. G. Skarmeta, M. Z. Izquierdo, and J. B. Blaya, "Neural networks for sonar and infrared sensors fusion," in *Proceedings of the Third International Conference on Information Fusion*, vol. 2, pp. 18–25, France, 10–13 July 2000.

[37] V. Genovese, E. Guglielmelli, A. Mantuano, G. Ratti, A. M. Sabatini, and P. Dario, "Low-cost, redundant proximity sensor system for spatial sensing and color-perception," *Electronics Letters*, vol. 31, pp. 632–633, 13 April 1995.

[38] A. M. Sabatini, V. Genovese, E. Guglielmelli, A. Mantuano, G. Ratti, and P. Dario, "A low-cost composite sensor array combining ultrasonic and infrared proximity sensors," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, pp. 120–126, Pittsburgh, PA, U.S.A., 5–9 August 1995.

[39] B. Chen and J. K. Tugnait, "Multisensor tracking of a maneuvering target in clutter using IMMPDA fixed-lag smoothing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, pp. 983–991, July 2000.

[40] Y. M. Chen and H. C. Huang, "Fuzzy logic approach to multisensor data association," *Mathematics and Computers in Simulation*, vol. 52, pp. 399–412, 15 July 2000.

[41] P. M. Novotny and N. J. Ferrier, "Using infrared sensors and the Phong illumination model to measure distances," in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 2, pp. 1644–1649, Detroit, MI, U.S.A., 10–15 May 1999.

[42] B. Andò and S. Graziani, "A new IR displacement system based on noise added theory," in *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference*, vol. 1, pp. 482–485, Budapest, Hungary, 21–23 May 2001.

[43] L. Korba, S. Elgazzar, and T. Welch, "Active infrared sensors for mobile robots," *IEEE Transactions on Instrumentation and Measurement*, vol. 43, pp. 283–287, April 1994.

[44] K. Hashimoto, T. Tsuruta, K. Morinaka, and N. Yoshiike, "High performance human information sensor," *Sensors and Actuators A—Physical*, vol. 79, pp. 46–52, 25 January 2000.

[45] N. Yoshiike, K. Morinaka, K. Hashimoto, M. Kawaguri, and S. Tanaka, "360 degrees direction type human information sensor," *Sensors and Actuators A—Physical*, vol. 77, pp. 199–208, 2 November 1999.

[46] P. J. de Groot, G. J. Postma, W. J. Melssen, and L. M. C. Buydens, "Validation of remote, on-line, near-infrared measurements for the classification of demolition waste," *Analytica Chimica Acta*, vol. 453, pp. 117–124, 18 February 2002.

[47] D. M. Scott, "A 2-color near-infrared sensor for sorting recycled plastic waste," *Measurement Science and Technology*, vol. 6, pp. 156–159, February 1995.

[48] G. Benet, F. Blanes, J. E. Simó, and P. Pérez, "Using infrared sensors for distance measurement in mobile robots," *Robotics and Autonomous Systems*, vol. 40, pp. 255–266, 30 September 2002.

[49] F. J. Cuevas, M. Servin, and R. Rodriguez-Vera, "Depth object recovery using radial basis functions," *Optics Communications*, vol. 163, pp. 270–277, 15 May 1999.

[50] P. Klysubun, G. Indebetouw, T. Kim, and T. C. Poon, "Accuracy of three-dimensional remote target location using scanning holographic correlation," *Optics Communications*, vol. 184, pp. 357–366, 15 October 2000.

[51] J. J. Esteve-Taboada, P. Refregier, J. Garcia, and C. Ferreira, "Target localization in the three-dimensional space by wavelength mixing," *Optics Communications*, vol. 202, pp. 69–79, 1 February 2002.

[52] B. Iske, B. Jäger, and U. Rückert, "A ray-tracing approach for simulating recognition abilities of active infrared sensor arrays," *IEEE Sensors Journal*, vol. 4, pp. 237–247, April 2004.

[53] Ç. Yüzbaşıoğlu and B. Barshan, "Improved range estimation using simple infrared sensors without prior knowledge of surface characteristics," *Measurement Science and Technology*, vol. 16, pp. 13905–1409, July 2005.

[54] H. V. Christensen, *Position Detection Based on Intensities of Reflected Infrared Light*. PhD thesis, Aalborg University, Department of Control Engineering, Denmark, November 2005.

[55] B. Hassibi and D. G. Stork, "Second-order derivatives for network pruning: optimal brain surgeon," in *Advances in Neural Information Processing Systems* (S. J. Hanson, J. D. Cowan, and C. L. Giles, eds.), vol. 5, pp. 164–171, Morgan Kaufmann, San Mateo, CA, 1993.

[56] T. Aytaç and B. Barshan, "Rule-based target differentiation and position estimation based on infrared intensity measurements," *Optical Engineering*, vol. 42, pp. 1766–1771, June 2003.

[57] T. Aytaç and B. Barshan, "Differentiation and localization of targets using infrared sensors," *Optics Communications*, vol. 210, pp. 25–35, September 2002.

[58] B. Barshan and T. Aytaç, "Position-invariant surface recognition and localization using infrared sensors," *Optical Engineering*, vol. 42, pp. 3589–3594, December 2003.

[59] T. Aytaç and B. Barshan, "Simultaneous extraction of geometry and surface properties of targets using simple infrared sensors," *Optical Engineering*, vol. 43, pp. 2437–2447, October 2004.

[60] T. Aytaç and B. Barshan, "Surface differentiation by parametric modeling of infrared intensity scans," *Optical Engineering*, vol. 44, pp. 067202–1–9, June 2005.

[61] B. Barshan, T. Aytaç, and Ç. Yüzbaşıoğlu, "Target differentiation with simple infrared sensors using statistical pattern recognition techniques," *Pattern Recognition*. accepted for publication, January 2007.

[62] Matrix Elektronik, AG, Kirchweg 24 CH-5422 Oberehrendingen, Switzerland, *IRS-U-4A Proximity Switch Datasheet*, 1995.

[63] M. D. Adams, "Lidar design, use, and calibration concepts for correct environmental detection," *IEEE Transactions on Robotics and Automation*, vol. 16, pp. 753–761, December 2000.

[64] A. Foessel-Bunting, "Radar sensor model for three-dimensional map building," in *Proceedings SPIE: Mobile Robots XV and Telemanipulator and*

*Telepresence Technologies VII* (H. M. Choset, D. W. Gage, and M. R. Stein, ed.), vol. 4195, pp. 127–138, November 2000, published March 2001.

[65] Arrick Robotics, P.O. Box 1574, Hurst, Texas, 76053 URL: www.robotics.com/rt12.html, *RT-12 Rotary Positioning Table*, 2002.

[66] T. Aytaç, "Differentiation and localization using infrared sensors," Master's thesis, Bilkent University, Department of Electrical Engineering, Ankara, Turkey, August 2002.

[67] B. Barshan and R. Kuc, "Differentiating sonar reflections from corners and planes by employing an intelligent sensor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 560–569, June 1990.

[68] B. Ayrulu and B. Barshan, "Identification of target primitives with multiple decision-making sonars using evidential reasoning," *International Journal of Robotics Research*, vol. 17, pp. 598–623, June 1998.

[69] J. W. Goodman, *Introduction to Fourier Optics*, pp. 246–249. New York: McGraw-Hill, 2nd ed., 1996.

[70] B. C. Bai and N. H. Farhat, "Learning networks for extrapolation and radar target identification," *Neural Networks*, vol. 5, pp. 507–529, May/June 1992.

[71] M. Cohen, H. Franco, N. Morgan, D. Rumelhart, and V. Abrash, "Context-dependent multiple distribution phonetic modelling with MLPs," in *Advances in Neural Information Processing Systems* (S. J. Hanson, J. D. Cowan, and C. L. Giles, ed.), pp. 649–657, San Mateo, CA: Morgan Kaufmann, 1993.

[72] K. S. Narendra and K. Parthasarathy, "Gradient methods for the optimization of dynamic systems containing neural networks," *IEEE Transactions on Neural Networks*, vol. 2, pp. 252–262, March 1991.

[73] M. I. Jordan and R. A. Jacobs, "Learning to control an unstable system with forward modeling," in *Advances in Neural Information Processing Systems 2* (D. S. Touretzky, ed.), pp. 324–331, San Mateo, CA: Morgan Kaufmann, 1990.

[74] M. Galicki, H. Witte, J. Dörschel, M. Eiselt, and G. Griessbach, "Common optimization of adaptive processing units and a neural network during the learning period: application in EEG pattern recognition," *Neural Networks*, vol. 10, pp. 1153–1163, August 1997.

[75] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Handwritten digit recognition with a back-propagation network," in *Advances in Neural Information Processing Systems 2* (D. S. Touretzky, ed.), pp. 396–404, San Mateo, CA: Morgan Kaufmann, 1990.

[76] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagation errors," *Nature*, vol. 323, pp. 533–536, 1986. London.

[77] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: John Wiley & Sons, 2001.

[78] M. Nørgaard, O. Ravn, N. K. Poulsen, and L. K. Hansen, *Neural Networks for Modelling and Control of Dynamic Systems*. London: Springer-Verlag, 2000.

[79] *Neural Network based Systems Identification Toolbox*. URL: http://www.iau.dtu.dk/research/control/nnsysid.html.

[80] S. K. Nayar, K. Ikeuchi, and T. Kanade, "Surface reflection: Physical and geometrical perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 611–634, July 1991.

[81] E. R. Davies, *Machine Vision: Theory, Algorithms, Practicalities*. London: Academic Press, 1990.

[82] M. Born and E. Wolf, *Principles of Optics*. Oxford, U.K.: Pergamon Press, 1980.

[83] B. T. Phong, "Illumination for computer generated pictures," *Communications of the ACM*, vol. 18, pp. 311–317, June 1975.

[84] M. A. Garcia and A. Solanas, "Automatic distance measurement and material characterization with infrared sensors," in *Proceedings of the 8th RoboCup International Symposium*, Lisbon, Portugal, 4–5 July 2004.

[85] M. A. Garcia and A. Solanas, "Estimation of distance to planar surfaces and type of material with infrared sensors," in *Proceedings of the 17th International Conference on Pattern Recognition*, vol. 1, pp. 745–748, Cambridge, U.K., 23–26 August 2004.

[86] G. Petryk and M. Buehler, "Dynamic object localization via a proximity sensor network," in *Proceedings IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 337–341, Washington D.C., 8–11 December 1996.

[87] T. Coleman, M. A. Branch, and A. Grace, *MATLAB Optimization Toolbox, User's Guide*. 1999.

[88] R. P. W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. de Ridder, and D. M. J. Tax, *A Matlab Toolbox for Pattern Recognition, PRTools4*. 2004.

[89] *PRTools toolbox*. URL: http://www.prtools.org/.

[90] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832–837, 1956.

[91] V. K. Rohatgi, *An Introduction to Probability Theory and Mathematical Statistics*. New York: John Wiley & Sons, 1976.

[92] G. J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*. New York: John Wiley & Sons, 1992.

[93] J. Bilmes, "A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models," tech. rep., University of Berkeley, ICSI-TR-97-021, 1997.

[94] A. Webb, *Statistical Pattern Recognition*. New York: John Wiley & Sons, 2002.

[95] E. Fix and J. L. Hodges, "Discriminatory analysis, nonparametric discrimination, consistency properties," Tech. Rep. 21-49-004 4, United States Air Force, School of Aviation Medicine, Randolph Field, Texas, 1951.

[96] M. P. Wand and M. C. Jones, *Kernel Smoothing*. London: Chapman and Hall, 1995.

[97] S. J. Sheather, "The performance of six popular bandwidth selection methods on some real data sets," *Computational Statistics*, vol. 7, pp. 225–250, 1992.

[98] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1986.

[99] S. Haykin, *Neural Networks: A Comprehensive Foundation*. New Jersey: Prentice Hall, 1994.

[100] V. N. Vapnik, *Estimation of Dependences Based on Empirical Data*. [in Russian], Moscow: Nauka, 1979. (English translation: New York: Springer Verlag, 1982).

[101] V. N. Vapnik, *Statistical Learning Theory*. Wiley-Interscience, 1998.

[102] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," *The Journal of Machine Learning Research*, vol. 2, pp. 45–66, 2002.