

EFFECTS OF ATMOSPHERIC CORRECTION ON VEHICLE
CLASSIFICATION WITH SINGLE AND DUAL BAND INFRARED IMAGES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

SEÇKİN ÖZSARAÇ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

AUGUST 2014

Approval of the thesis:

**EFFECTS OF ATMOSPHERIC CORRECTION ON VEHICLE
CLASSIFICATION WITH SINGLE AND DUAL BAND INFRARED IMAGES**

submitted by **SEÇKİN ÖZSARAÇ** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Gönül Turhan Sayan
Head of Department, **Electrical and Electronics Engineering**

Prof. Dr. Gözde Bozdağı Akar
Supervisor, **Electrical and Electronics Eng. Dept., METU**

Examining Committee Members:

Prof. Dr. A. Aydın Alatan
Electrical and Electronics Engineering Dept., METU

Prof. Dr. Gözde Bozdağı Akar
Electrical and Electronics Engineering Dept., METU

Assoc. Prof. Dr. İlkey Ulusoy
Electrical and Electronics Engineering Dept., METU

Assoc. Prof. Dr. Alptekin Temizel
Graduate School of Informatics, METU

Assist. Prof. Dr. Behçet Uğur Töreyn
Electrical and Electronics Eng. Dept., Çankaya University

Date:

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: SEÇKİN ÖZSARAÇ

Signature :

ABSTRACT

EFFECTS OF ATMOSPHERIC CORRECTION ON VEHICLE CLASSIFICATION WITH SINGLE AND DUAL BAND INFRARED IMAGES

ÖZSARAÇ, SEÇKİN

Ph.D., Department of Electrical and Electronics Engineering

Supervisor : Prof. Dr. Gözde Bozdağı Akar

August 2014, 155 pages

A vehicle classification system, which uses features based on radiometry, is developed for both single band and dual band infrared (IR) image sequences. For classification using dual band sequences, the process is divided into six components. These are registration, fusion, moving vehicle detection, geometry estimation, atmospheric effects removal, and classification. In the single band case, registration and fusion steps are not used. The first major novelty of the thesis is an atmospheric correction, i.e. atmospheric effects removal, system that considers the spectral characteristics of the detector, lens, and filter. In this system, an enhanced temperature calibration method is developed and it is shown that the temperature accuracy for the dynamic range of the IR camera is very close to the ultimate goal, i.e. Noise Equivalent Temperature Difference (NETD) value of the camera. Furthermore, as the atmospheric effects vary from pixel to pixel, a geometry estimation method is developed to estimate the Line Of Sight (LOS) geometry for each pixel using only the Global Positioning System (GPS) coordinates of the camera and a Point Of Interest (POI) in the scene. The second major novelty of the thesis lies in the usage of the atmospherically corrected radiance values as features to improve the classification performance of the detected objects. The motivation is, each vehicle class has a discriminating radiance value that originates from the source temperature of the object modified by the intrinsic characteristics of the radiating surface. As a consequence, significant performance gains are

achieved due to the use of radiance values in classification both for a single band and a dual band measurement systems.

Keywords: Atmospheric Correction, Radiometric Camera Calibration, Line of Sight Estimation, Vehicle Classification, Moving Vehicle Detection, IR Image Fusion, IR Image Registration

ÖZ

TEK VE ÇİFT BANT KIZILÖTESİ İMGELERLE ARAÇ SINIFLANDIRMAYA ATMOSFERİK DÜZELTMENİN ETKİLERİ

ÖZSARAÇ, SEÇKİN

Doktora, Elektrik ve Elektronik Mühendisliği Bölümü

Tez Yöneticisi : Prof. Dr. Gözde Bozdağı Akar

Ağustos 2014 , 155 sayfa

Tek bant ve çift bant kızılötesi (KÖ) imge dizileri için, ışınım ölçümüne dayalı öznelikler kullanan, bir araç sınıflandırma sistemi geliştirilmiştir. Süreç, çift bant dizilerinde sınıflandırma için, altı bileşene bölünmüştür. Bunlar çakıştırma, kaynaştırma, hareketli araç algılama, geometri kestirme, atmosferik etkileri giderme ve sınıflandırmadır. Tek bant durumunda, çakıştırma ve kaynaştırma adımları kullanılmamaktadır. Tezin getirdiği birinci ana yenilik, sezimci, lens ve süzgecin izgesel karakteristiklerinin hesaba katıldığı, atmosferik düzeltme yani atmosferik etki giderme sistemidir. Bu sistemde, iyileştirilmiş bir sıcaklık kalibrasyon yöntemi geliştirilmiştir ve KÖ kameranın dinamik çalışma aralığı için yöntemin sıcaklık doğruluğunun nihai hedefe yani Gürültü Eşdeğer Sıcaklık Farkına (NETD) çok yakın olduğu gösterilmiştir. Ayrıca, pikselden piksele atmosfer etkilerinin değişmesi sebebiyle, sadece kameranın Küresel Konumlandırma Sistemi (GPS) ve sahnedeki bir İlgi Noktası (POI) koordinatlarını kullanarak her pikselin Görüş Hattı (LOS) geometrisini kestirmek için bir geometri kestirim yöntemi geliştirilmiştir. Tezin getirdiği ikinci ana yenilik, algılanmış nesnelerin sınıflandırma başarımını artırmak için atmosfer etkilerinden arındırılmış ışınırılık değerlerinin öznitelik olarak kullanılmasında yatmaktadır. Buradaki güdülenme; her araç sınıfının, nesnenin asıl sıcaklığının ışınım yapan yüzeyin için özellikleri ile değiştirilmesi sonucu oluşan, ayırıcı ışınırılık değerlerine sahip olmasıdır. Sonuç olarak, tek bant ve çift bant ölçüm sistemleri için sınıflandırmada ışınırılık değerlerinin kulla-

nılmasına baęlı olarak önemli başarımlar kazanılmıřtır.

Anahtar Kelimeler: Atmosferik Düzeltme, Iřınım ile İlgili Kamera Kalibrasyonu, Görüş Hattı Kestirimi, Araç Sınıflandırma, Hareketli Araç Algılama, KÖ İmge Kaynaştırma, KÖ İmge Çakıştırma

To Nihan

ACKNOWLEDGMENTS

I gratefully thank my supervisor Prof. Dr. Gözde Bozdağı Akar for her endless and sincere support, constructive criticism, encouragement, and supervision throughout the development of the thesis. I greatly appreciate her patience and belief on me and also her guidance related to my academic life.

I express my gratitude to thesis monitoring committee members Assoc. Prof. Dr. İlkey Ulusoy and Assoc. Prof. Dr. Alptekin Temizel. They have closely monitored the progress of the study and the quality of the thesis has been remarkably improved with their valuable comments and fruitful discussions with them.

I also would like to thank thesis jury members Prof. Dr. A. Aydın Alatan and Assist. Prof. Dr. B. Uğur Töreyn for sparing their time and effort in the evaluation of the thesis.

I thank my employer Tübitak Bilgem İltaren and also my colleagues there for all the support on the study.

I offer my best appreciation to my parents E. Aysel Öz Saraç and Kahraman Öz Saraç, and also to my brothers İlker Öz Saraç and Ömer Öz Saraç for their continuous support which kept me up in any circumstances. Without them, everything were impossible.

Lastly, and most importantly, I wish to thank my beloved wife Nihan Öz Saraç and my dearest son Emir Can Öz Saraç. They have always supported me and shared anything with me. To them, I dedicate this thesis.

TABLE OF CONTENTS

ABSTRACT	v
ÖZ	vii
ACKNOWLEDGMENTS	x
TABLE OF CONTENTS	xi
LIST OF TABLES	xviii
LIST OF FIGURES	xix
LIST OF ABBREVIATIONS	xxiii
CHAPTERS	
1 INTRODUCTION	1
1.1 Scope of the Thesis	1
1.2 Outline of the Thesis	4
2 REGISTRATION	7
2.1 Affine Registration	7
2.2 Perspective Registration	8
2.3 Utilized Registration Method	9
3 ATMOSPHERIC CORRECTION SYSTEM	11

3.1	Geometry Estimation	14
3.2	Atmospheric Effects Removal	17
3.2.1	Object Radiance Estimation	20
3.2.2	Radiative Transfer Computations	23
3.2.2.1	Visibility Estimation	24
3.2.3	Temperature Calibration with Radiometric Inter- polation	24
3.2.3.1	Practical Issues	26
4	DATA FUSION	27
4.1	Symbol Level (Information) Fusion	28
4.2	Object Level (Decision) Fusion	28
4.3	Feature Level Fusion	28
4.3.1	Feature Selection Based Fusion	29
4.3.2	Feature Extraction Based Fusion	29
4.4	Pixel Level (Image) Fusion	29
4.4.1	Space Domain Based Image Fusion	30
4.4.1.1	Pixel Averaging Based Image Fusion	30
4.4.1.2	Max-Min Method Based Image Fusion	30
4.4.1.3	Color Composite (False Color) Based Image Fusion	31
4.4.1.4	PCA Method Based Image Fusion	31
4.4.1.5	Entropy Based Image Fusion	31

4.4.2	Transform Domain Based Image Fusion	32
4.4.2.1	DCT Based Image Fusion	32
4.4.3	Multiscale Decomposition Domain Based Image Fusion	33
4.4.3.1	Pyramid Transform Based Image Fusion	33
4.4.3.2	Wavelet Transform Based Image Fusion	33
4.5	Utilized Data Fusion Method	35
5	MOVING VEHICLE DETECTION	37
5.1	Background Subtraction	37
5.1.1	Frame Differencing	37
5.1.2	Moving Average Filtering	38
5.1.3	Single Gaussian	38
5.1.4	Mixture of Gaussians	39
5.1.5	Non-parametric Approach	42
5.1.6	GraphCut	44
5.1.7	GrabCut	45
5.1.8	GrayCut	46
5.1.9	Object-based Segmentation	47
5.2	Utilized Background Subtraction Method	47
5.3	Post-processing After Background Subtraction	47
5.3.1	Erosion	48

5.3.2	Dilation	48
5.3.3	Closing	49
5.3.4	Opening	49
5.3.5	Connected Component Labeling	49
5.4	Utilized Post-processing After Background Subtraction Method	50
6	CLASSIFICATION	53
6.1	Feature Extraction	57
6.1.1	Blob Features	58
6.1.2	Radiometric Features	59
6.2	Classifiers	60
6.2.1	Nearest Neighbor Classifier	60
6.2.2	k-nearest Neighbor Classifier	60
6.2.3	Support Vector Machines	61
6.2.3.1	Multi-class SVM	62
6.3	Utilized Classification Method	63
6.4	Performance Metrics	63
6.4.1	Confusion Matrix	64
6.4.2	Precision	64
6.4.3	Recall	64
6.4.4	Accuracy	65
7	MEASUREMENT SETUP	67

7.1	Single Band Measurement Setup	67
7.2	Dual Band Measurement Setup	67
8	ANALYSIS RESULTS	69
8.1	Registration Results	69
8.2	Geometry Estimation Results	70
8.2.1	Geometry Estimation Results for the Single Band Measurements	70
8.2.2	Geometry Estimation Results for the Dual Band Measurements	72
8.3	Atmospheric Effects Removal Results	74
8.3.1	Radiative Transfer Computations Results	75
8.3.2	Temperature Calibration with Radiometric Inter- polation Results	76
8.3.2.1	Temperature Calibration with Radio- metric Interpolation Results for the Sin- gle Band Measurements	76
8.3.2.2	Temperature Calibration with Radio- metric Interpolation Results for the Dual Band Measurements	77
8.3.3	Atmospheric Effects Removal Results for the Sin- gle Band Measurements	79
8.3.4	Atmospheric Effects Removal Results for the Dual Band Measurements	86
8.4	Data Fusion Results	97
8.5	Moving Vehicle Detection Results	98

8.5.1	Moving Vehicle Detection Results for the Single Band Measurements	99
8.5.2	Moving Vehicle Detection Results for the Dual Band Measurements	100
8.6	Classification Results	102
8.6.1	Effects of Radiometric Calibration on Vehicle Classification	104
8.6.2	Effects of Atmospheric Correction on Vehicle Classification	108
8.6.3	Classification Results for the Single Band Measurements	112
8.6.4	Classification Results for the Dual Band Measurements	116
8.6.5	Effects of Occlusion on Vehicle Classification	120
9	CONCLUSION	123
	REFERENCES	125
	APPENDICES	
A	INFRARED THEORY	139
A.1	Planck's Radiation Law	139
A.2	Wien's Displacement Law	140
A.3	Stefan-Boltzman Law	140
B	INFRARED DETECTORS	143
B.1	Quantum Well Detectors	143
B.2	Photo Voltage Detectors	144

B.3	Micro Bolometer Detectors	144
C	GEOCENTRIC EARTH RADIUS	145
D	MULTIRESOLUTION THEORY	147
D.1	Wavelet Transform	147
E	INFRARED SIGNATURE ANALYSIS SOFTWARE	149
	CURRICULUM VITAE	153

LIST OF TABLES

TABLES

Table 6.1	Confusion Matrix Example	64
Table 8.1	Classification with DN Features Confusion Matrix	106
Table 8.2	Classification with DN Features Performance Metrics	106
Table 8.3	Classification with Radiometric Features Confusion Matrix	107
Table 8.4	Classification with Radiometric Features Performance Metrics	107
Table 8.5	Classification with Blackbody Radiance Features Confusion Matrix	110
Table 8.6	Classification with Blackbody Radiance Features Performance Metrics	110
Table 8.7	Single Band Measurement Setup Reference Method Confusion Matrix	113
Table 8.8	Single Band Measurement Setup Reference Method Performance Metrics	113
Table 8.9	Single Band Measurement Setup Proposed Method Confusion Matrix	114
Table 8.10	Single Band Measurement Setup Proposed Method Performance Metrics	114
Table 8.11	Dual Band Measurement Setup Reference Method Confusion Matrix	117
Table 8.12	Dual Band Measurement Setup Reference Method Performance Metrics	118
Table 8.13	Dual Band Measurement Setup Proposed Method Confusion Matrix	118
Table 8.14	Dual Band Measurement Setup Proposed Method Performance Metrics	118

LIST OF FIGURES

FIGURES

Figure 1.1	Reference Method for the Single Band Measurement System	2
Figure 1.2	Proposed Method for the Single Band Measurement System	3
Figure 1.3	Reference Method for the Dual Band Measurement System	4
Figure 1.4	Proposed Method for the Dual Band Measurement System	5
Figure 3.1	Flow Chart of the Atmospheric Correction System	13
Figure 3.2	Coordinate System Definitions	15
Figure 3.3	Flow Chart of the Atmospheric Effects Removal Method	18
Figure 3.4	Flow Chart of the Radiative Transfer Computations	23
Figure 5.1	A Sample Result for the Moving Vehicle Detection Method	50
Figure 5.2	A Sample Result for the Utilized Post-processing After Background Subtraction Method	51
Figure 7.1	Spectral Transfer Functions of the Dual Band Measurement Setup .	68
Figure 8.1	Dual Band Measurement Setup Registration Results - Near Scenario	70
Figure 8.2	Dual Band Measurement Setup Registration Results - Far Scenario	71
Figure 8.3	Single Band Measurement Setup Geometry Estimation Results - Near Scenario	72
Figure 8.4	Single Band Measurement Setup Geometry Estimation Results - Far Scenario	72
Figure 8.5	Dual Band Measurement Setup Geometry Estimation Results - Near Scenario	73

Figure 8.6 Dual Band Measurement Setup Geometry Estimation Results - Far Scenario	74
Figure 8.7 Effect of the Range on the Radiative Transfer	75
Figure 8.8 Effect of the Zenith Angle on the Radiative Transfer	75
Figure 8.9 Single Band Measurement Setup - Justification of the Temperature Calibration Method	76
Figure 8.10 Single Band Measurement Setup - Radiometric Calibration	77
Figure 8.11 Dual Band Measurement Setup - Justification of the Temperature Calibration Method	78
Figure 8.12 Dual Band Measurement Setup - Radiometric Calibration	80
Figure 8.13 Single Band Measurement Setup Atmospheric Effects Removal Analysis Results - Near Scenario	81
Figure 8.14 Single Band Measurement Setup Atmospheric Effects Removal Radiance Images - Near Scenario	82
Figure 8.15 Single Band Measurement Setup Atmospheric Effects Removal Analysis Results - Far Scenario	83
Figure 8.16 Single Band Measurement Setup Atmospheric Effects Removal Radiance Images - Far Scenario	84
Figure 8.17 Single Band Measurement Setup Atmospheric Effects Removal Results	85
Figure 8.18 Spectral Reflectance of the Asphalt for the Single Band Measurement Setup	85
Figure 8.19 Dual Band Measurement Setup Atmospheric Effects Removal Analysis Results - MWIR Camera - Near Scenario	86
Figure 8.20 Dual Band Measurement Setup Atmospheric Effects Removal Radiance Images - MWIR Camera - Near Scenario	87
Figure 8.21 Dual Band Measurement Setup Atmospheric Effects Removal Analysis Results - LWIR Camera - Near Scenario	88
Figure 8.22 Dual Band Measurement Setup Atmospheric Effects Removal Radiance Images - LWIR Camera - Near Scenario	89
Figure 8.23 Dual Band Measurement Setup Atmospheric Effects Removal Analysis Results - MWIR Camera - Far Scenario	90

Figure 8.24 Dual Band Measurement Setup Atmospheric Effects Removal Radiance Images - MWIR Camera - Far Scenario	91
Figure 8.25 Dual Band Measurement Setup Atmospheric Effects Removal Analysis Results - LWIR Camera - Far Scenario	92
Figure 8.26 Dual Band Measurement Setup Atmospheric Effects Removal Radiance Images - LWIR Camera - Far Scenario	93
Figure 8.27 Dual Band Measurement Setup Atmospheric Effects Removal Results - MWIR Camera	94
Figure 8.28 Dual Band Measurement Setup Atmospheric Effects Removal Results - LWIR Camera	95
Figure 8.29 Spectral Reflectance of the Asphalt for the Dual Band Measurement Setup	96
Figure 8.30 Air Temperature Measurements for the Dual Band Measurement Setup	96
Figure 8.31 Dual Band Measurement Setup Data Fusion Results - Near Scenario	98
Figure 8.32 Dual Band Measurement Setup Data Fusion Results - Far Scenario	99
Figure 8.33 Single Band Measurement Setup Moving Vehicle Detection Results - Near Scenario	100
Figure 8.34 Single Band Measurement Setup Moving Vehicle Detection Results - Far Scenario	100
Figure 8.35 Dual Band Measurement Setup Moving Vehicle Detection Results - Near Scenario	101
Figure 8.36 Dual Band Measurement Setup Moving Vehicle Detection Results - Far Scenario	101
Figure 8.37 Single Band Measurement Setup <i>DN</i> Feature Extraction Results .	104
Figure 8.38 Single Band Measurement Setup Radiometric Feature Extraction Results	105
Figure 8.39 Single Band Measurement Setup <i>DN</i> vs Radiometric Feature Comparison for the Van Class	105
Figure 8.40 Classification with <i>DN</i> vs. Radiometric Features Cross Validation Results	106

Figure 8.41 Single Band Measurement Setup Blackbody Feature Extraction Results	109
Figure 8.42 Classification with Blackbody Radiance vs. Radiometric Features Cross Validation Results	109
Figure 8.43 Single Band Measurement Setup Blob Feature Extraction Results	112
Figure 8.44 Single Band Measurement Setup Cross Validation Results	113
Figure 8.45 Single Band Measurement Setup Misclassification Examples	115
Figure 8.46 Dual Band Measurement Setup Blob Feature Extraction Results	116
Figure 8.47 Dual Band Measurement Setup Radiometric Feature Extraction Results	117
Figure 8.48 Dual Band Measurement Setup Cross Validation Results	117
Figure 8.49 Dual Band Measurement Setup Misclassification Examples	119
Figure E.1 GUI of IRSA	150
Figure E.2 GUI of Language Selection	150
Figure E.3 GUI of IRSA in Turkish	150
Figure E.4 GUI of IRSA - Apparent Temperature Selection	151
Figure E.5 GUI of IRSA - Segmented Image Selection	152

LIST OF ABBREVIATIONS

2D	Two Dimensional
3D	Three Dimensional
AIRS	Alliance Icing Research Study
ATR	Automatic Target Recognition
BG	Background
BRDF	Bidirectional Reflectance Distribution Function
CNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
DN	Digital Number
DOG	Difference Of Gaussians
EM	Expectation Maximization
ETM+	Enhanced Thematic Mapper Plus
FAST	Features from Accelerated Segment Test
FG	ForeGround
FOV	Field Of View
FRAM	Fog Remote sensing And Modeling
GaAs	Gallium Arsenide
GKLT	Generalized Karhunen Loève Transform
GLOH	Gradient Location and Orientation Histogram
GMM	Gaussian Mixture Model
GPS	Global Positioning System
GUI	Graphical User Interface
HgCdTe	Mercury Cadmium Telluride
HH	High-High
HL	High-Low
HOG	Histogram of Oriented Gradients

ICA	Independent Component Analysis
InSb	Indium antimonide
IR	InfraRed
IRSA	InfraRed Signature Analysis software
IT	Integration Time
LDA	Linear Discriminant Analysis
LH	Low-High
LL	Low-Low
LOG	Laplacian Of Gaussians
LOS	Line Of Sight
LVQ	Learning Vector Quantization
LWIR	Long Wave InfraRed
MAP	Maximum A Posteriori
MB	Micro Bolometer
MCMF	Minimum-Cut/Maximum-Flow
ML	Maximum Likelihood
MLP	MultiLayer Perceptron
MNN	Modular Neural Network
MOG	Mixture Of Gaussians
MPM	Multinomial Pattern Matching
MSER	Maximally Stable Extremal Regions
MWIR	Middle Wave InfraRed
NETD	Noise Equivalent Temperature Difference
NIR	Near InfraRed
NN	Nearest Neighbor
NUC	Non-Uniformity Correction
ORB	Oriented FAST and Rotated BRIEF
PCA	Principle Component Analysis
PD	Probability of Detection
PFA	Probability of False Alarm
POI	Point Of Interest
PV	Photo Voltage
QWIP	Quantum Well Infrared Photodetector
RBF	Radial Basis Function

RGB	Red Green Blue
RH	Relative Humidity
ROI	Region Of Interest
RT	Radiative Transfer
RX	Reed-Xiaoli
RUC	Rapid Update Cycle
SIFT	Scale Invariant Feature Transformation
SPRT	Sequential Probability Ratio Test
SURF	Speeded Up Robust Features
SVM	Support Vector Machines
TISIE	Temperature Independent Spectral Indices of Emissivity
TM	Thematic Mapper
VQ	Vector Quantizer

CHAPTER 1

INTRODUCTION

Infrared (IR) imaging equipment are widely used in various applications such as security, traffic control, surveillance, defense industry, automotive industry, building diagnostics, automation, gas detection, electrical defect inspection, mechanical defect inspection, and also as test and radiometric measurement tools.

For security, traffic control, and surveillance purposes; vehicle classification is an important application area. Classification using videos captured with visible band cameras is a well studied topic [1–21]. Besides, vehicle classification in IR image sequences is proposed in [22–36]. However, to author’s knowledge there is no study that uses the radiometry related information in vehicle classification. In this context, a methodology is proposed to be used with either a single band IR camera system or a dual band IR camera system. The novelty of the method is the usage of the radiometric quantities in the classification. Moreover, to make the features geometry independent, atmospheric effects removal¹ method including the IR camera’s radiometric calibration is proposed, too.

1.1 Scope of the Thesis

In this thesis, two methods for vehicle classification is proposed. The first one is for the single band IR camera case. In order to evaluate the performance of the proposed method a reference system will be used and its flow chart is given in Fig.1.1. In this

¹ Throughout the thesis, “atmospheric effects removal” and “atmospheric correction” terms are used interchangeably.

setup, the digital data output of the IR camera is used as input. Digital data, which is the raw output of the IR camera, is called Digital Number (DN) values. DN data is fed into a vehicle detection block so that the objects of interests are separated from the background (BG) and vehicle blobs are obtained. Hence, the Region Of Interests (ROI) for each blob are given at the output. ROI data is then fed into classification process and the process outputs the class label of the object. The classes are limited to vehicle types car, van, truck/bus, and BG clutter. In this setup, the only manual operation is the selection of a line on the frame so that only a single vehicle blob intersects with this line on the frame.

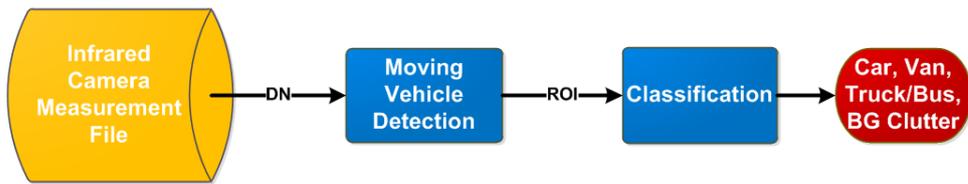


Figure 1.1: Reference Method for the Single Band Measurement System

To enhance the classification performance; the radiometric quantities, e.g. temperature, radiance, etc., corresponding to the ROI data are proposed to be used in the classification process. The flow chart of the proposed method for the single band measurement case is shown in Fig.1.2. The raw output of the IR camera, that is DN , contains only the contrast information about the scene. DN values do not contain any radiometric information if they are used alone. On the other hand, if DN data is used together with a radiometric calibration process, DN data can be converted to radiometric quantities. However, according to [37–40] these quantities strongly depend on the atmospheric propagation conditions especially when the distance between the IR camera and the ROI is large. Therefore, DN data and the prior information are firstly used to find the geometry between the camera and the scene. Prior information comprises Global Positioning System (GPS) of the camera, GPS of a predetermined point in the scene, and also the FOV value of the camera. Then, the geometry data that is the altitude (h), zenith angle (θ), and range (R) values are used in the atmospheric effects removal process. At the output of the relative process, geometry independent data that is called source radiance data is computed. Finally, the source radiance data is used together with the ROI data in the classification process. Hence, a performance increase in the classification is expected since the source radiance contains discrimi-

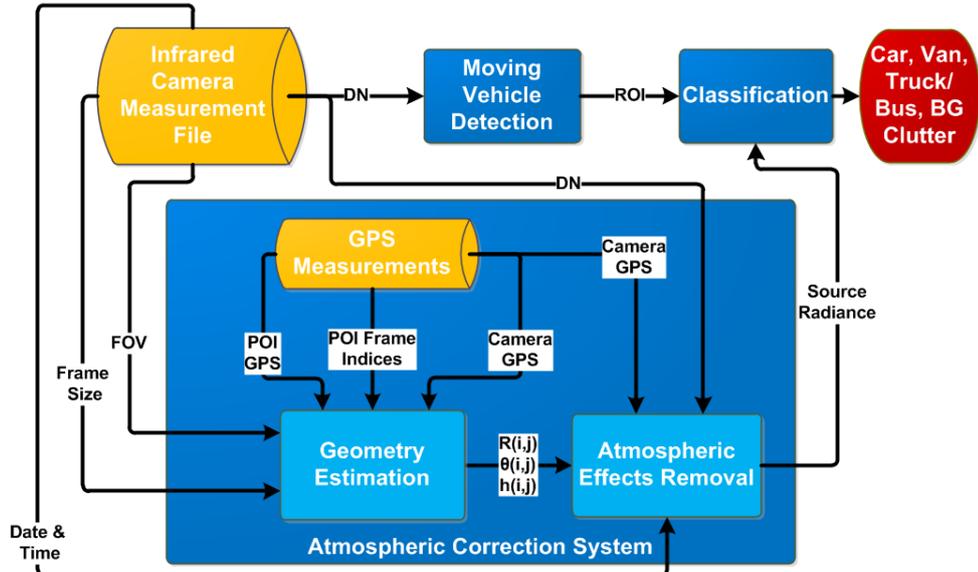


Figure 1.2: Proposed Method for the Single Band Measurement System

natory features for each class. For example, the apparent temperature of the wheel of a truck is relatively higher than the temperature of the wheel of a car. Therefore, the source radiance distribution of a wheel of the object of interest is a valuable feature for classification.

The second method is proposed for a dual channel system. The system may be either a single camera responsive to two different IR bands or two different IR cameras. The flow chart of the reference method for the dual band system is shown in Fig.1.3. In this setup, DN data of the Middle Wave IR (MWIR) band and DN data of the Long Wave IR (LWIR) band are used. Firstly, $MWIR DN$ and $LWIR DN$ are registered for the case of two different cameras. Then the registered data are fed into the fusion process. At the output of the fusion process, a single $Fused DN$ data is obtained, which is an enhanced image in terms of contrast. The rest of the system is the same with the single band reference system.

Similar to the single band case, the radiometric quantities of dual band are proposed to be used in the classification process. The flow chart of the proposed method for the dual band system is shown in Fig.1.4. The difference between the reference system is, the registered DN data is used to find the h , R , and θ data for each channel. Actually, these data are the same for each channel but they are drawn separately for completeness since atmospheric effects removal process uses these data together with the DN data of the respective channels. Hence, geometry independent $MWIR$

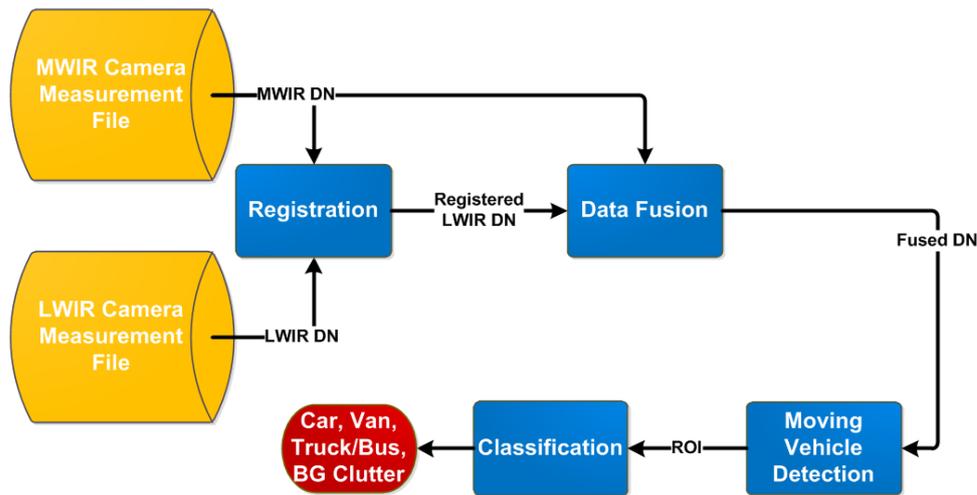


Figure 1.3: Reference Method for the Dual Band Measurement System

Source Radiance and *LWIR Source Radiance* are estimated and given as input to the classification process. Finally, the source radiance data of two bands are used in the classification process together with the *ROI* data of the vehicle detection process. Hence, a performance increase in the classification is expected similar to the single band case. In the dual band case, a better performance gain is expected since the source radiance of two different IR bands shall be more discriminative than the single band case. This is because, according to Wien's law hotter objects shall yield larger contrast with respect to the background in the MWIR band, whereas colder objects shall yield larger contrast in the LWIR band.

1.2 Outline of the Thesis

In Chapter 2 the registration process, which is used only for the dual band measurement system, is given. Chapter 3 explains the details of the proposed atmospheric correction system, which contains the geometry estimation and atmospheric effects removal methods. In Chapter 4, data fusion, which is again used only for the dual band measurement system, is explained. Literature review and also the utilized fusion method are given in this chapter, too. Moving vehicle detection method is explained in Chapter 5 again with a detailed literature review. The final stage of the system is the classification method and it is given in Chapter 6 together with all the related studies on vehicle classification. Chapter 7 presents both measurement setups used in the thesis. In Chapter 8, the analysis results of each method are given and Chapter 9

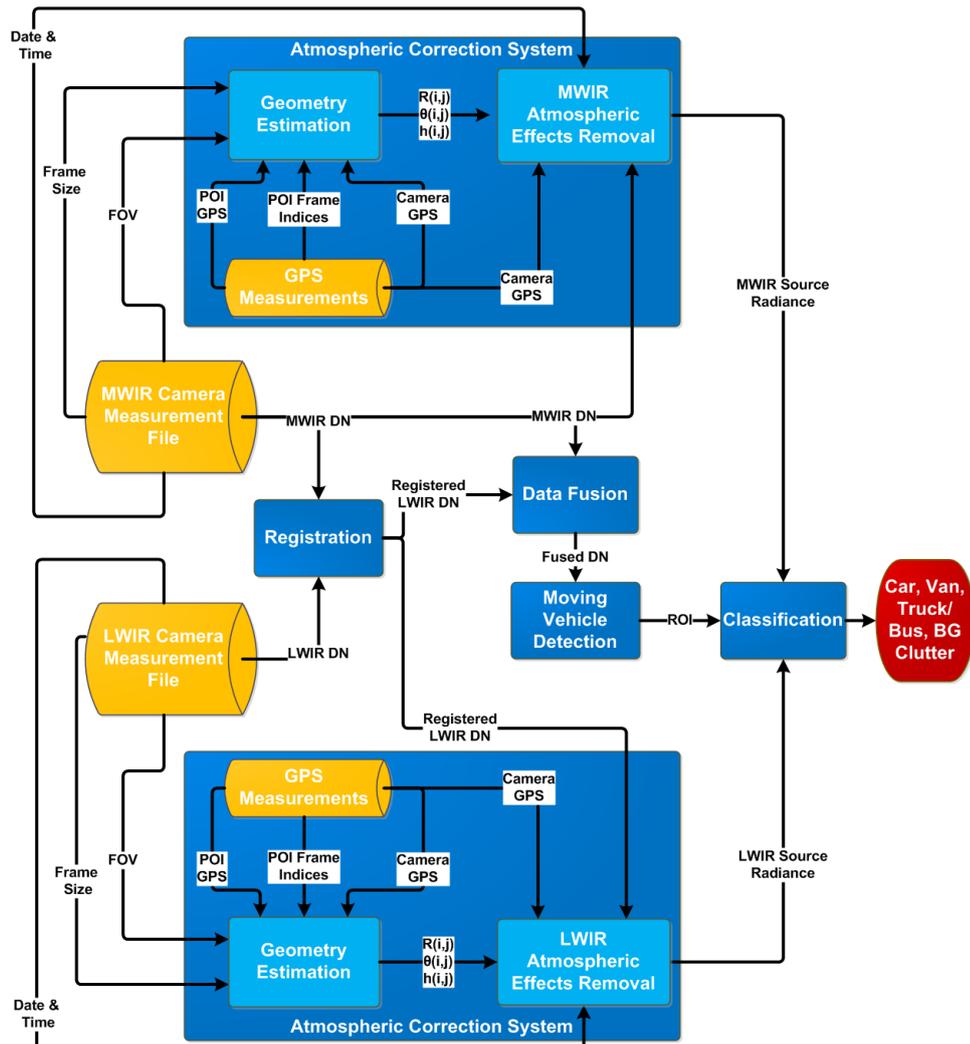


Figure 1.4: Proposed Method for the Dual Band Measurement System

concludes the thesis with a brief summary and potential future work.

In Appendix A, the basics of the infrared theory is given. Infrared theory is used in the derivation of the atmospheric effects removal method. In Appendix B, the detector types are explained so that the characteristics of the IR cameras of measurement setups can be tracked. Appendix C explains the geocentric earth radius used in the geometry estimation method derivations. A brief review of multiresolution theory, which is used in data fusion process, is given in Appendix D. The implemented software, which is called Infrared Signature Analysis software (IRSA), is explained in Appendix E. IRSA is used for starting from loading the IR image sequences until to the end of feature extraction. Hence, all the processes in the flow chart of the system except the classifier are handled by IRSA.

CHAPTER 2

REGISTRATION

When multiple IR detectors are used, captured frames do not represent the same projection of the scene. Therefore, in the applications where the projection of the same scene with various sensors is desired, registration process is used.

2.1 Affine Registration

In [41], IR images of a patient are captured. Then, using the special markers on the body, patient movement is compensated using affine transformation for temporal analysis purposes. The captured digital photo potentially yields a rotated, scaled, shifted, and slightly sheared image of the area of the interest. The equation of the affine transformation is,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \quad (2.1)$$

where x and y are the image coordinates of the original frame. Primed coordinates are the coordinates obtained by image registration. a_{ij} 's for $i, j = 1, 2$ are the parameters of the affine transformation that is to be estimated. a_{13} and a_{23} are the bias coordinates. To estimate a_{ij} values, at least three non-collinear landmarks are needed.

2.2 Perspective Registration

In [42], the perspective registration is given by,

$$x' = \frac{a_{11}x + a_{12}y + a_{13}}{b_{11}x + b_{12}y + 1} \quad (2.2)$$

$$y' = \frac{a_{21}x + a_{22}y + a_{23}}{b_{21}x + b_{22}y + 1} \quad (2.3)$$

where a_{ij} and b_{kl} for $i, k, l = 1, 2$ and $j = 1, 2, 3$ are the parameters of the perspective registration. Clearly, if $b_{kl} = 0 \forall k, l$ then the registration becomes an affine registration. In the proposal, the authors approximate the perspective transformation as piecewise affine. The reference image is divided into tiles and an affine registration is applied in order to find the best matching tiles in the registered image. Then, the estimated optimum tiles are used to compute the global perspective transformation. The parameters are computed iteratively using a variation of Levenberg-Marquardt non-linear least squares optimization method. While the parameters of affine registration is being evaluated, the search criterion is to minimize the sum of squared distances between the two images. However, in the multi-band IR case, the pixel values are not expected to be the same even for the same object (See Sec.3.2). In addition, images corresponding to different IR bands clearly have dissimilar contrast. Therefore, the mentioned perspective registration method is not suitable for our case.

In [43], a hierarchical image registration algorithm is proposed. Log-polar transform is used in the spatial domain rather than in the frequency domain to cope with similarity transformation consisting of rotation, scale, and translation. Then, the log-polar transform is coupled with a non-linear least squares algorithm to estimate the perspective transformation parameters. The log-polar registration step aims to align the two images using only rotation, scale, and translation. Hence, a fine initial estimate for the following perspective registration step is obtained. Perspective registration step, which is based on non-linear least squares optimization with Levenberg-Marquardt algorithm, yields sub-pixel accuracy. However, the optimization problem is basically the minimization of the difference between the pixel values of the reference and the registered images. The assumption of the method is that there is an equality or a linear relationship between the pixel values of the two images. As a result, similar to the previous work the proposed perspective registration method is not applicable to our

case due to the constraints of the dual band IR imaging.

2.3 Utilized Registration Method

In this thesis, a controlled measurement setup has been used. A fair amount of rotation, translation, scaling, and shearing is expected. Rotation exists because two cameras are mounted on two separate tripods. A small translation is expected as two different cameras are used. Although theoretical field of view (FOV) values of the cameras are the same, they are different in reality and therefore there is a small amount of scaling issue. Finally, as the center of rotation is not the same, shearing is expected as well. Since all of these effects are minor, the linear affine transformation is suitable for the registration process of the dual band case. Hence, a set of manually chosen points in the MWIR band image that correspond to a set of points in the LWIR band image are used for the estimation of transformation.

Adapting the method in [44], Eq.(2.1) is rewritten as,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.4)$$

With n number of manually chosen landmark points, the following linear system of equations is obtained,

$$\begin{bmatrix} x'_1 & \dots & x'_n \\ y'_1 & \dots & y'_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \\ 1 & \dots & 1 \end{bmatrix} \quad (2.5)$$

where x_i and y_i are the image coordinates of the i^{th} sample in original frame. Primed coordinates are the ones of the registered frame.

Let the following matrix definitions be made as,

$$\mathcal{X}' \triangleq \begin{bmatrix} x'_1 & \dots & x'_n \\ y'_1 & \dots & y'_n \end{bmatrix} \quad (2.6)$$

$$\mathcal{A} \triangleq \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad (2.7)$$

$$\mathcal{X} \triangleq \begin{bmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \\ 1 & \dots & 1 \end{bmatrix} \quad (2.8)$$

Then Eq.(2.5) becomes,

$$\mathcal{X}' = \mathcal{A}\mathcal{X} \quad (2.9)$$

The affine transform matrix can be estimated using the pseudo-inverse technique as,

$$\hat{\mathcal{A}} = \mathcal{X}'\mathcal{X}^T (\mathcal{X}\mathcal{X}^T)^{-1} \quad (2.10)$$

where $\hat{\mathcal{A}}$ is the least-squares estimate of the affine transform matrix, \mathcal{A} .

CHAPTER 3

ATMOSPHERIC CORRECTION SYSTEM

The captured data with IR detectors is affected by the atmospheric variations. Accurate compensation of atmospheric effects of data captured by an IR camera is crucial for several applications such as vegetation monitoring, temperature monitoring, satellite images, hyperspectral imaging, simulation and modeling, surface properties characterization, and IR measurement interpretation. Atmospheric effects depend on the temporal changes, i.e. year, season, day, hour, etc., and on the Line Of Sight (LOS) geometry between the camera and the measured scene. The orientation and the optical depth of the camera significantly affect the variation of the geometry across the pixels. Therefore, the removal of the atmospheric effects is an important task in the above mentioned applications where IR data independent of temporal and spatial variations is desired.

In [45], atmospheric compensation for satellite images in MWIR band is given, which is based on user defined parameters. Hence, the method lacks physical modeling. In [46], a physical approach to correct the atmospheric effects is proposed. Using Modtran[®] [47] with the measured meteorological parameters and the temperature of a specific point in the scene, the surface brightness values for different times of a day are estimated. For the LOS geometry, high resolution digital surface model (DSM) is utilized. However, DSM is available only for a limited number of regions on the world.

There are a number of atmospheric correction studies for satellite images. In [48], an atmospheric correction method based on calibrated radiance measurements is given. The meteorological parameters are measured and also the Temperature Independent

Spectral Indices of Emissivity (TISIE) is computed using the night time field measurements. Day time directional reflectivity values are computed using the TISIE data. A Bidirectional Reflectance Distribution Function (BRDF) model is obtained from these directional reflectivity values. Then, by combining all these data atmospheric correction is applied with the help of Modtran[®] and the surface brightness temperature values are obtained. Since the method uses night time measurements to obtain the BRDF model, it is impractical for applications where a field measurement is not possible. In [49], an atmospheric correction method based on meteorological measurements is proposed for the specific Landsat-5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) bands. Since the method depends on field measurements of the object being measured, it is again impractical. A method based on vicarious calibration, which uses the output of the radiometric calibration process as input, is given in [50] and the method yields atmospherically corrected radiance values. The method is specialized to estimate the water-leaving radiance values in the Near IR (NIR) and visible bands. Another atmospheric correction method that is specific to Landsat 7 ETM+ band and depends on field measurements, is proposed in [51]. The method uses the normalized spectral response of the sensor, meteorological measurements and also the emissivity measurements of the object of interest.

A novel atmospheric correction system is developed. The main advantage of the proposed system is the accuracy due to the physical measurements and also the detailed consideration of the radiometric relationships. On the other hand, the method does require specific GPS and meteorological measurements as opposed to the user defined modeling approach. Our atmospheric effects removal method is based on the Radiative Transfer (RT) components, i.e. spectral atmospheric transmittance and path radiance with scattering, for the specific LOS geometry of each pixel. However, there is no DSM with an appropriate resolution for our measurement site in Ankara. Therefore, as the first contribution, we propose a geometry estimation method in order to be able to accurately compute the atmospheric profile. The geometry of each pixel for the measured scene is computed using the proposed process, where only two GPS coordinates are used. While obtaining the range and zenith angle of each pixel, the shape of the earth is considered and the view vector of the camera is also found.

Then the geometry data and the measured meteorological data are used to obtain the spectral transmittance and path radiance with Modtran5[®]. Unlike most of the atmospheric correction studies present in the literature, our method does not rely on any measurement or knowledge of the intrinsic properties of the scene such as emissivity, BRDF, temperature, etc. This is because, our primary goal is to output a radiometric quantity, which is independent of temporal and spatial variations as much as possible. In addition, we aim to find a quantity, which requires low computation power. Hence, we give a detailed derivation of the atmospheric correction with the radiometric relationships that considers the spectral characteristics of the detector, lens, and filter. In order to have a complete system, we introduce an enhanced temperature calibration method for the IR cameras that reaches a temperature accuracy very close to the ultimate goal, i.e. Noise Equivalent Temperature Difference (NETD) value of the camera.

The proposed system mainly consists of two methods, which are the geometry estimation and the atmospheric effects removal. Geometry estimation aims to find the range and zenith angle values corresponding to each pixel together with the altitude of the camera. Then, the output of the geometry estimation is fed into atmospheric effects removal process. These variables are used to compute the pixel-wise atmospheric transmittance and path radiance with scattering. Afterwards, these RT components are removed from the apparent radiance values, which are obtained with the help of proposed radiometric calibration method. The flow chart of the proposed system is shown in Fig.3.1.

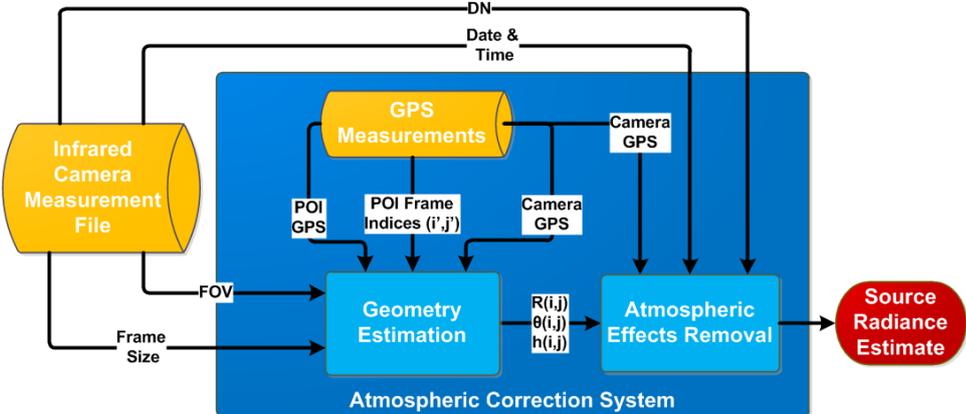


Figure 3.1: Flow Chart of the Atmospheric Correction System

3.1 Geometry Estimation

To estimate the geometry between the camera and the points corresponding to each pixel on a frame, it is assumed that two GPS data are available. These are the GPS coordinate of the camera and the GPS coordinate of a Point Of Interest (POI). POI is a predetermined point in the scene that is projected to a pixel on the frame.

In practice, almost all the GPS measurement devices output geodetic latitude rather than geocentric latitude. Hence, we assume that there are two sets of geodetic latitude, longitude, and altitude measurements. Let these values be lat_{cam} , lon_{cam} , and h_{cam} for the GPS of the camera and lat_{POI} , lon_{POI} , and h_{POI} for the POI. In addition, it is assumed that the horizontal and vertical FOV, FOV_h and FOV_v , of the camera and also the horizontal and vertical frame sizes, m and n , are known. The camera is located at a higher location than the measured scene, that is $h_{cam} > h_{POI}$ and also it is directed downwards.

In order to derive the geometric equations, a coordinate system shall be defined. Let the origin of the coordinate system be at the core of the earth, y be the axis that is parallel to the tangential plane at the camera location, z be the axis increasing upwards with respect to the ground and x be the corresponding perpendicular axis. Coordinate system axes together with the image plane, angle, and distance definitions are shown in Fig.3.2(a), Fig.3.2(b), and Fig.3.2(c).

To find the range and zenith angle for each pixel, the radius of the earth shall be used. The earth is not a perfect sphere so effective radius (R_{eff}) of the earth or in other words the radius of curvature at zero altitude shall be obtained. The effective earth radius, whose derivation is given in App.C, is as follows

$$R_{eff} = \sqrt{\frac{R_{eq}^4 \cos(lat_{cam})^2 + R_{pol}^4 \sin(lat_{cam})^2}{R_{eq}^2 \cos(lat_{cam})^2 + R_{pol}^2 \sin(lat_{cam})^2}} \quad (3.1)$$

where R_{eq} is the radius of the earth at the equator and R_{pol} is the radius at the pole.

Then, altitudes for each pixel shall be determined. In this context, it is assumed that the points corresponding to each pixel on the frame lie on the same altitude with the altitude of the POI, that is to say all the points are isoplanar. This assumption holds as long as the scene is a nearly flat surface. However, structures such as buildings,

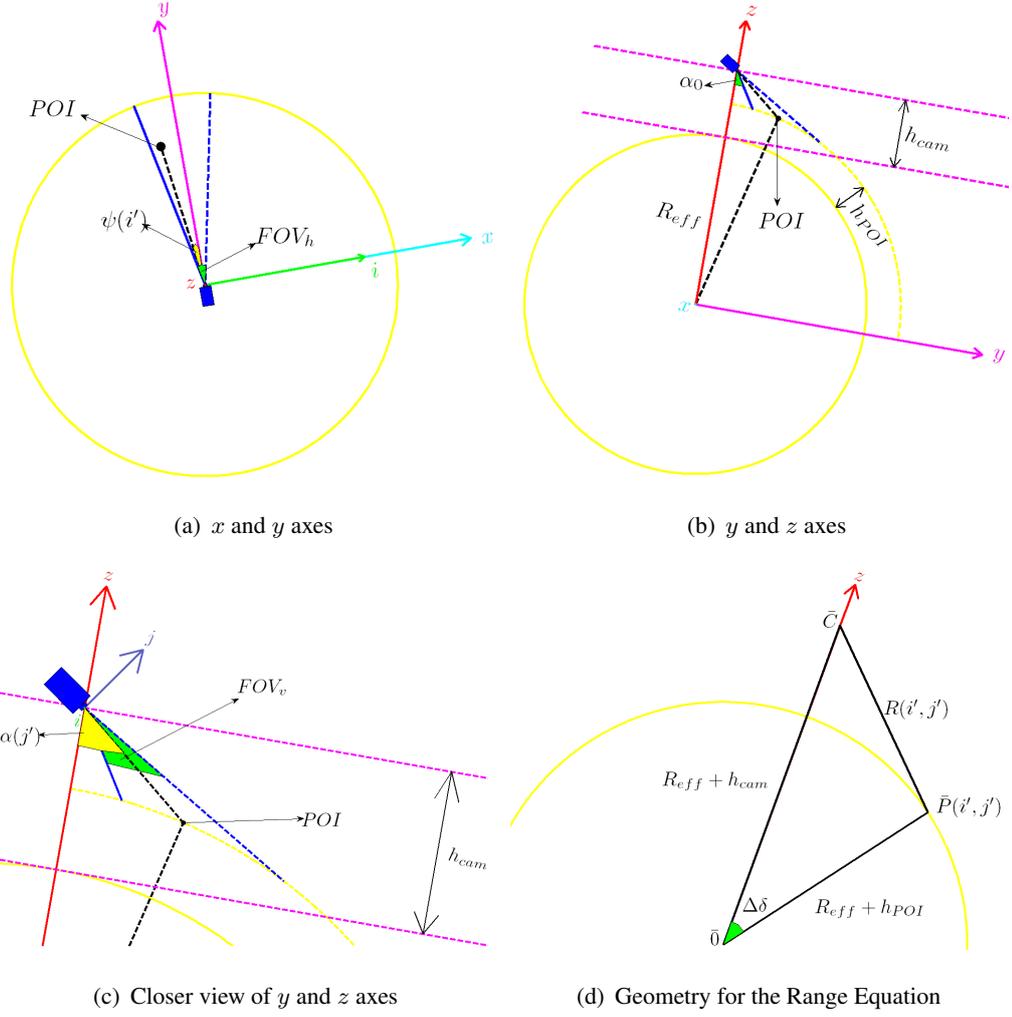


Figure 3.2: Coordinate System Definitions

hills, valleys, etc. decrease the accuracy of the proposed method's output for the corresponding pixels.

On an isoplanar surface, the altitude of a pixel, $h(i, j)$, is given by

$$h(i, j) = h_{POI} \quad (3.2)$$

where i and j are the respective horizontal and vertical frame coordinates on the image plane. Thus, the scene is isoplanar with a radius of $R_{eff} + h_{POI}$. Since the optical depth of IR cameras is on the order of a few kilometers it can be assumed that the isoplanar surface in the FOV of the camera lies on a sphere. Therefore, the equality for a sphere is written as follows,

$$x^2 + y^2 + z^2 = (R_{eff} + h_{POI})^2 \quad (3.3)$$

Afterwards, the azimuth angle is simply written as,

$$\psi(i, j) = \psi(i) = \arctan\left(\frac{x}{y}\right) \quad \forall j = 1, 2, \dots, n \quad (3.4)$$

The elevation angle is written as,

$$\alpha(i, j) = \alpha(j) = \arctan\left(\frac{y}{R_{eff} + h_{cam} - z}\right) \quad \forall i = 1, 2, \dots, m \quad (3.5)$$

Combining Eqs.(3.3-3.5), we get the following equality.

$$y^2 \tan^2 \psi(i) + y^2 + [R_{eff} + h_{cam} - y \cot \alpha(j)]^2 = (R_{eff} + h_{POI})^2 \quad (3.6)$$

$$y^2 \left[1 + \tan[\psi(i)]^2 + \frac{1}{\tan[\alpha(j)]^2} \right] - 2 \cot[\alpha(j)](R_{eff} + h_{cam})y + h_{cam}^2 - h_{POI}^2 + 2R_{eff}(h_{cam} - h_{POI}) = 0 \quad (3.7)$$

Let y_r be the proper root of Eq.(3.7) and let the following vector definitions are made

$$\bar{C} \triangleq [0, 0, R_{eff} + h_{cam}] \quad (3.8)$$

$$\bar{P}(i, j) \triangleq [y_r \tan \psi(i), y_r, R_{eff} + h_{cam} - y_r \cot \alpha(j)] \quad (3.9)$$

where \bar{C} and $\bar{P}(i, j)$ are row vectors that hold the three dimensional (3D) coordinates of the camera and the 3D coordinates of the projected point (i, j) , respectively. Then, the range between the camera and a point in the scene can be written as,

$$R(i, j) = \|\bar{C} - \bar{P}(i, j)\|_2 \quad (3.10)$$

and the zenith angle is computed as,

$$\theta(i, j) = \pi - \arccos\left[\frac{h_{cam} - h_{POI}}{R(i, j)}\right] \quad (3.11)$$

Up to this point, azimuth and elevation angles are defined using the 3D coordinates of the points in the scene as shown in Fig.3.2. Alternatively, they can be easily computed using the two dimensional (2D) image plane coordinates, i and j , and also the respective FOV value as,

$$\psi(i) = \left(i - \frac{m}{2} - \frac{1}{2}\right) \frac{FOV_h}{m} \quad (3.12)$$

$$\alpha(j) = \left(j - \frac{1}{2}\right) \frac{FOV_v}{n} + \alpha_0 \quad (3.13)$$

for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$ where α_0 is the angle between the lower limit of the FOV_v and the z axis as shown in Fig.3.2(b).

The only unknown in Eq.(3.12) and Eq.(3.13) is α_0 value, which can be found by using the GPS coordinates of the camera and the POI in the FOV. Assume the POI is projected to the indices (i', j') on the image plane, where the geometry is shown in Fig.3.2(d). Then, the central angle between the two GPS coordinates is found by [52]

$$\Delta\delta = \arccos[\sin lat_{cam} \sin lat_{POI} + \cos lat_{cam} \cos lat_{POI} \cos(lon_{cam} - lon_{POI})] \quad (3.14)$$

and the range value for the POI is found using the cosine theorem with the following equation.

$$R(i', j') = [(R_{eff} + h_{POI})^2 + (R_{eff} + h_{cam})^2 - 2(R_{eff} + h_{POI})(R_{eff} + h_{cam}) \cos \Delta\delta]^{0.5} \quad (3.15)$$

Once $R(i', j')$ is found, we can write the elevation angle of the POI using the geometrical identities as

$$\alpha(j') = \arctan \left[\frac{R(i', j') \sin \theta(i', j') \cos \psi(i')}{h_{cam}} \right] \quad (3.16)$$

Then, α_0 is computed using Eq.(3.11), Eq.(3.12), Eq.(3.16), and Eq.(3.13) in the given order after replacing i and j with the POI indices i' and j' . At the end, since all the unknowns are found, $\bar{P}(i, j)$ vector can be calculated $\forall i, j$ using Eq.(3.12), Eq.(3.13), Eq.(3.6) and Eq.(3.9) in order. Then, using Eq.(3.10) and Eq.(3.11) range and zenith angle values are estimated for each pixel.

3.2 Atmospheric Effects Removal

Once the LOS geometry between the camera and the measured scene is found, the atmospheric correction method computation can be done. Our goal in this method is to come up with a radiometric quantity, in which the influence of the atmosphere is subtracted. Moreover, we seek a quantity, which is possible to compute with a very low computational complexity. Consequently, we aim to compute the radiance of

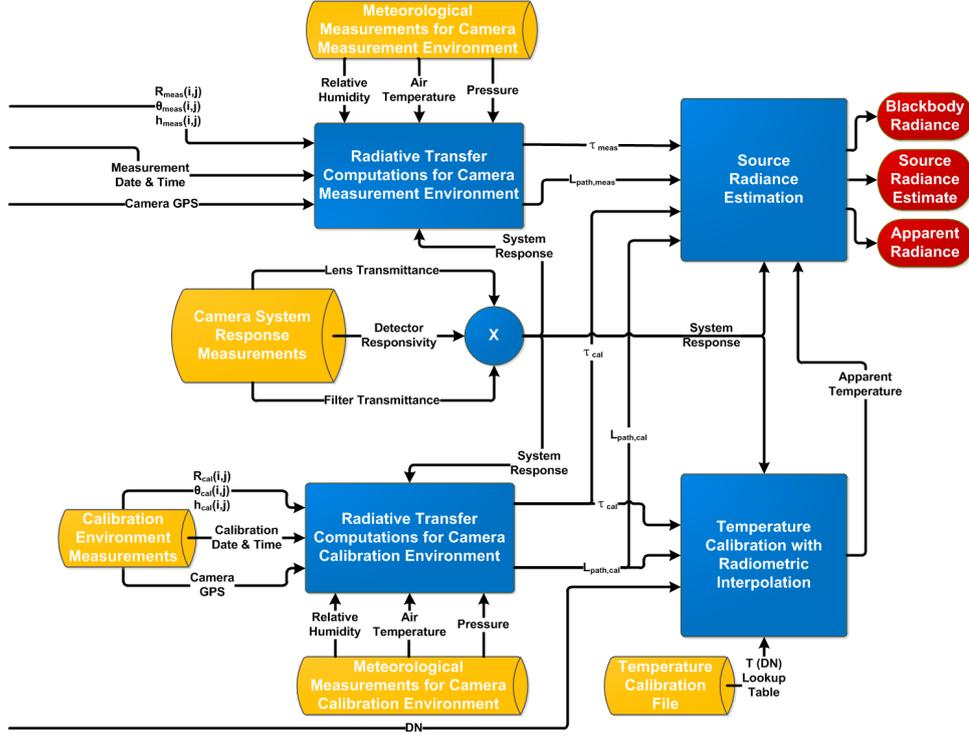


Figure 3.3: Flow Chart of the Atmospheric Effects Removal Method

the object of interest without atmospheric effects. The flow chart of the atmospheric effects removal method is shown in Fig.3.3.

To start with, the well-known image irradiance equation [53] is written as,

$$E(i, j) = L(i, j) \frac{\pi}{4} \left(\frac{d}{f} \right)^2 \cos^4 \phi(i, j) \quad (3.17)$$

where E is the irradiance [Wm^{-2}], L is the scene radiance [$Wsr^{-1}m^{-2}$], d is the aperture diameter of the imaging system [mm], f is the focal length of the imaging system [mm] and ϕ is the angle subtended by the principal ray from the optical axis [rad].

In [54, 55], the brightness of the image that the camera records is given with the relation,

$$I(i, j) = E(i, j)t_{exp} \quad (3.18)$$

where I is the image brightness and t_{exp} is the time that the detector is exposed to the scene radiation, i.e. Integration Time (IT).

IR cameras measure the apparent pass-band integrated spectral scene radiance. Hence,

Eq.(3.17) and Eq.(3.18) shall be written as a function of wavelength,

$$E_x(i, j, \lambda) = L_x(i, j, \lambda) \frac{\pi}{4} \left(\frac{d}{f} \right)^2 \cos^4 \phi(i, j) \quad (3.19)$$

where λ represents the wavelength [μm], E_x is the spectral irradiance [$W m^{-2} \mu m^{-1}$], and L_x is the spectral radiance [$W sr^{-1} m^{-2} \mu m^{-1}$]. Subscript x denotes the component being recorded, a car, a calibration source, etc.

$$I_x(i, j) = \int_{\lambda_{low}}^{\lambda_{high}} E_x(i, j, \lambda) t_{exp} \varphi_{det}(i, j, \lambda) \varphi_{lens}(i, j, \lambda) \varphi_{filter}(i, j, \lambda) d\lambda \quad (3.20)$$

where φ_{det} is the normalized spectral responsivity of the IR camera detector, φ_{lens} is the spectral transmittance of the IR camera lens, φ_{filter} is the spectral transmittance of the filter used in the IR camera, λ_{low} and λ_{high} are the respective lower and upper wavelength limits that the responsivity of the IR camera is greater than zero.

Adapting from [46], the apparent spectral scene radiance can be computed with the following relationship,

$$L_x(i, j, \lambda) = L_{emis,x}(i, j, \lambda) + L_{atm,x}(i, j, \lambda) + L_{ref,x}(i, j, \lambda) \quad (3.21)$$

where $L_{emis,x}$ is the self-emission of the recorded object, $L_{atm,x}$ is the path radiance component in the recorded geometry, and $L_{ref,x}$ is the reflected radiance from the object being recorded. $L_{emis,x}$ is given by

$$L_{emis,x}(i, j, \lambda) \triangleq L_{bb}[\lambda, T_x(i, j)] \epsilon_x(i, j, \lambda) \tau_x[\lambda, R_x(i, j), h_x(i, j), \theta_x(i, j)] \quad (3.22)$$

T_x is the surface temperature [K] of the object being recorded, ϵ_x is the spectral emissivity of the object, τ_x is the spectral atmospheric transmittance between the object and the IR camera depending on the range, altitude, and zenith angle geometry that are R_x , h_x , and θ_x respectively. L_{bb} is the spectral blackbody radiance [$W sr^{-1} m^{-2} \mu m^{-1}$] computed using Planck's Law given in Eq.(A.2).

$L_{atm,x}$ can be computed using

$$L_{atm,x}(i, j, \lambda) \triangleq L_{path,x}[\lambda, R_x(i, j), h_x(i, j), \theta_x(i, j)] \quad (3.23)$$

where $L_{path,x}$ is the spectral atmospheric path radiance with scattering between the object and the IR camera depending on R_x , h_x , and θ_x for the environment condition during recordings.

For simplicity and feasibility, we assume $L_{ref,x}(i, j, \lambda) = 0$ for every $i \in [1, m]$, $j \in [1, n]$, and $\lambda \in [\lambda_{low}, \lambda_{high}]$. This assumption holds as long as the measurements are planned for the conditions with minimum reflections and the pass-band of the camera is close to the thermal band rather than the NIR band.

As we have defined the building blocks of the radiometric relationships, we will continue with the derivation of the estimation of our objective quantity, i.e. object radiance in the following subsection. Afterwards, both the RT computations and the proposed temperature calibration, which are used in object radiance estimation, will be given.

3.2.1 Object Radiance Estimation

To derive the object radiance estimation method, firstly let the spectral transfer function of the camera (φ_{cam}) be defined as,

$$\varphi_{cam}(i, j, \lambda) \triangleq \varphi_{det}(i, j, \lambda)\varphi_{lens}(i, j, \lambda)\varphi_{filter}(i, j, \lambda) \quad (3.24)$$

and let the terms that do not depend on the incident radiance be defined as,

$$\xi(i, j) \triangleq \frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos^4 \phi(i, j)t_{exp} \quad (3.25)$$

With the above mentioned definitions, the brightness recorded by the IR camera can be rewritten as

$$I_x(i, j) = \int_{\lambda_{low}}^{\lambda_{high}} \xi(i, j)L_x(i, j, \lambda)\varphi_{cam}(i, j, \lambda)d\lambda \quad (3.26)$$

Infrared cameras give DN at the output with a typical resolution of 14 to 16 bits. According to [55], there is a nonlinear relationship between the digital number and the brightness values. Therefore,

$$DN_x(i, j) = f[I_x(i, j)] \quad (3.27)$$

where $f(\cdot)$ is a monotonic increasing non-linear function. According to [55], \cos^4 effect is negligible since most modern visible camera lenses are designed to compensate for this effect and provide a nearly constant mapping between radiance and irradiance. However, this is not the case for IR cameras. To compensate this effect

and also the highly non-uniform response characteristics of the detector elements, Non-Uniformity Correction (NUC) process is applied in the IR cameras.

The NUC process linearizes the relation between the DN_x and I_x values. Then, without loss of generality, we can assume that $f(\cdot)$ is one-to-one as long as the quantization errors are relatively small, which is the case for 14 to 16 bit resolution IR cameras. Therefore, function $f(\cdot)$ can be treated as a linear invertible function as long as an appropriate NUC is used.

When NUC is applied, \cos^4 effect in the image irradiance equation is removed and the spatial uniformity of the camera's transfer function is obtained due to linearization. Hence, Eq.(3.26) simplifies to

$$I_x(i, j) = \int_{\lambda_{low}}^{\lambda_{high}} \xi L_x(i, j, \lambda) \varphi_{cam}(\lambda) d\lambda \quad (3.28)$$

For the rest of the thesis, it is assumed that an appropriate NUC is applied and the IR camera works in the linear regime¹. Hence, the terms that become independent of the frame coordinates will be simply written without the i and j indices.

Since our aim is to remove the atmospheric effects, the output of the IR camera for the measurements that is $DN_{x=meas}$ shall be used. Firstly, the following equalities hold as long as the same ξ and φ_{cam} is used for both the radiometric calibration and the measurement. This is the case when the same camera, lens, filter, IT, and NUC table are used. As a result, the following equations are obtained.

$$DN_{x=meas}(i, j) = DN_{x=cal}(i, j) \quad (3.29)$$

$$f[I_{x=meas}(i, j)] = f[I_{x=cal}(i, j)] \quad (3.30)$$

$$I_{x=meas}(i, j) = I_{x=cal}(i, j) \quad (3.31)$$

$I_{x=meas}$ and $I_{x=cal}$ are computed using Eq.(3.28). Then using Eq.(3.21), we can sim-

¹ Linear regime is the interval, where DN_x changes linearly with I_x .

plify Eq.(3.31) as,

$$\begin{aligned}
& \int_{\lambda_{low}}^{\lambda_{high}} L_{emis,x=meas}(i, j, \lambda) \varphi_{cam}(\lambda) d\lambda + \\
& \int_{\lambda_{low}}^{\lambda_{high}} L_{atm,x=meas}(i, j, \lambda) \varphi_{cam}(\lambda) d\lambda = \\
& \int_{\lambda_{low}}^{\lambda_{high}} L_{emis,x=cal}(i, j, \lambda) \varphi_{cam}(\lambda) d\lambda + \\
& \int_{\lambda_{low}}^{\lambda_{high}} L_{atm,x=cal}(i, j, \lambda) \varphi_{cam}(\lambda) d\lambda
\end{aligned} \tag{3.32}$$

τ_x and $L_{path,x}$ components present in Eq.(3.32) can be computed with a RT model software, e.g. Modtran[®] [46, 48, 49, 51]. Then, the right-hand side of Eq.(3.32) can be computed since all the related parameters are known. So, the first term of the left-hand side, which is the only component related to the object being measured, is calculated.

If the spectral emissivity of the measured object is known or measured, the surface temperature of the measured object can be computed iteratively. However, $\epsilon_{x=meas}$ is usually unknown, in which case $T_{x=meas}$ is impossible to find. On the other hand, the product of the emissivity and the blackbody radiance at the temperature $T_{x=meas}$ is almost independent of the atmospheric conditions in the steady-state, which implies the product does not depend on the geometry. Hence, the object radiance, L_{obj} [$W sr^{-1} m^{-2}$], which is the quantity that we are aiming to estimate is defined by,

$$L_{obj}(i, j) \triangleq \int_{\lambda_{low}}^{\lambda_{high}} L_{bb}[\lambda, T_{x=meas}(i, j)] \epsilon_{x=meas}(i, j, \lambda) d\lambda \tag{3.33}$$

A former approach for a similar aim, which does not take the spectral transfer function of the camera into account is given in [56]. Similarly, the dependency of the atmospheric propagation on the altitude and zenith angle is not considered in the related work, too. A more adequate and advanced approach is proposed in this thesis. Once the first term of the left hand-hand side of Eq.(3.32) is computed, the estimate of L_{obj} that is called the source radiance estimate, L_{src} , is computed using,

$$L_{src}(i, j) = \frac{\int_{\lambda_{low}}^{\lambda_{high}} L_{emis,x=meas}(i, j, \lambda) \varphi_{cam}(\lambda) d\lambda}{\tau_w(i, j)} \tag{3.34}$$

where

$$\tau_w(i, j) \triangleq \frac{1}{\lambda_{high} - \lambda_{low}} \int_{\lambda_{low}}^{\lambda_{high}} \tau_{x=meas}[\lambda, R_{x=meas}(i, j), h_{x=meas}(i, j), \theta_{x=meas}(i, j)] \varphi_{cam}(\lambda) d\lambda \quad (3.35)$$

It should be noted that $L_{src} \approx L_{obj}$ and the dependency of L_{src} on the frame coordinates is only due to the object location in the scene. L_{src} does not depend on the geometry, e.g. $R_{x=meas}$, because the geometric dependencies and also the atmospheric effects together with the influence of the camera's transfer function are removed by division with the weighted transmittance component, τ_w .

3.2.2 Radiative Transfer Computations

In order to be able to compute the source radiance estimates, the output of a RT model output, i.e. spectral transmittance and path radiance with scattering, is needed. For this purpose, we used Modtran5[®] RT model in which, atmosphere layers are constructed with the sea level air temperature, Relative Humidity (RH), and the pressure measurements. In the aerosol model of the RT model, visibility value is critical but we do not have the means to measure visibility. Therefore, visibility values are estimated using the models given in Sec.3.2.2.1. The rest of the parameters are used as the default values of the related MidLatitude Winter and MidLatitude Summer models depending on the measurement date. The flow chart of the RT computations is shown in Fig.3.4.

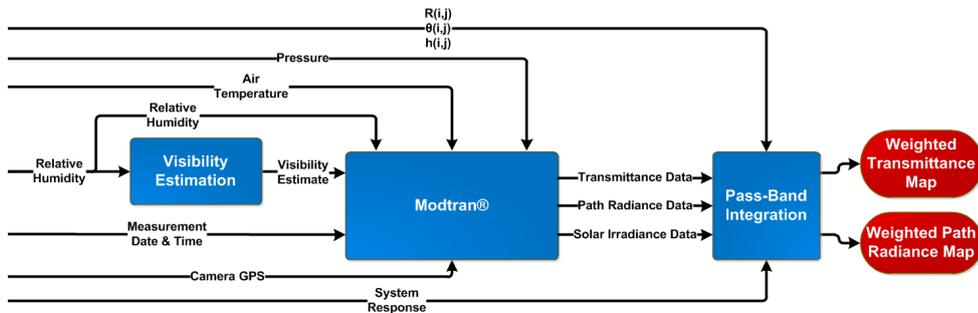


Figure 3.4: Flow Chart of the Radiative Transfer Computations

3.2.2.1 Visibility Estimation

According to [57], visibility depends on the temperature, humidity, solar zenith angle, total cloud cover, wind speed, and wind direction. Visibility (vis_{DMI}) is computed using these parameters so that only 2% of the light from the object is received at the visibility distance. The coefficients in the equations are optimized for Denmark and the dependency of the visibility on these parameters is non-linear.

In [58], visibility is computed using RH . There are three different models to compute the visibility. These are Rapid Update Cycle (RUC) model (vis_{RUC}), Alliance Icing Research Study (AIRS) model (vis_{AIRS}), and Fog Remote sensing And Modeling (FRAM) model (vis_{FRAM}). The relationships are given as,

$$vis_{RUC} = 60 \exp \left[-2.5 \frac{(RH - 15)}{80} \right] \quad (3.36)$$

$$vis_{AIRS} = -0.0177RH^2 + 1.462RH + 30.8 \quad (3.37)$$

$$vis_{FRAM} = -41.5 \ln(RH) + 192.3 \quad (3.38)$$

where RH denotes the relative humidity.

The former work is specialized to Denmark and so it is not suitable for our measurements in Ankara. Among the three models, AIRS model is valid only for a sub-interval of the RH range. Moreover, RUC model is the output of a relatively old study. Therefore, in this thesis we used the FRAM model to estimate the visibility values.

3.2.3 Temperature Calibration with Radiometric Interpolation

In order to give radiometric meaning to the DN values, calibration is required. Calibration is usually done either as temperature calibration [59] or as radiance calibration [51, 60–67]. Irradiance calibration, flux calibration [68], and also reflectance calibration [69–71] also exist in the literature. In this paper, an enhanced temperature calibration method, which takes into account the computed in-band apparent radiance values, is proposed. The aim of the proposed calibration is to match the DN values with the known surface temperatures of a calibration device.

The novelty of the proposed calibration method is that it considers the effect of the atmospheric environment in the temperature calibration process, emissivity of the calibration device, and also the spectral transfer function of the detector, lens, and the filter used in the camera. Although the calibration is done at very close distances, the spectral variation of the transmittance is effective due to CO_2 , ozone, and water vapor absorption. Therefore, the environment is taken into account while calculating the in-band apparent radiance values.

The temperature calibration outputs $T_{x=cal}$ for $DN_{x=cal}$ values in the dynamic range of the camera. But the related process yields temperature and DN pairs only for a limited number of calibration points due to practical limitations such as time and manpower constraints. A typical number of calibration points is 20, whereas the dynamic range of DN values due to the scene radiance in the linear working regime is on the order of 10^4 . Therefore, the temperature values for the remaining DN values, i.e. other than the calibration points, shall be obtained. A basic approach is to use linear interpolation for the temperature values corresponding to the in-between DN values. However, there is a non-linear relationship between the temperature and the IR camera measurements. For this reason, a method that interpolates the required values based on radiometric relationships is proposed. The core of the method is based on the requirement that the IR camera works in the linear regime and also an appropriate NUC process is applied. Once the related requirements are met, the image brightness changes linearly with DN . In this case, using Eq.(3.27) and Eq.(3.28) the following equality is written.

$$DN_{x=cal}(i, j) = f \left[\int_{\lambda_{low}}^{\lambda_{high}} \xi L_{x=cal}(i, j, \lambda) \varphi_{cam}(\lambda) d\lambda \right] \quad (3.39)$$

Since the temperature calibration is typically done at very close distances, the variation of the range, altitude, and zenith angle across the pixels is very small and so this variation has no practical effect on $\tau_{x=cal}$ and $L_{path,x=cal}$. In addition, the calibration device has an almost uniform surface so its surface temperature and spectral emissivity are assumed to be homogeneous and so independent of i and j . Therefore, without loss of generality the equation is simplified to

$$DN_{x=cal} = f [\xi L_{app,x=cal}(T_{x=cal})] \quad (3.40)$$

where

$$L_{app,x=cal}(T_{x=cal}) \triangleq \int_{\lambda_{low}}^{\lambda_{high}} [L_{bb}(\lambda, T_{x=cal})\epsilon_{x=cal}(\lambda) + \tau_{x=cal}(\lambda, R_{x=cal}, h_{x=cal}, \theta_{x=cal}) + L_{path,x=cal}(\lambda, R_{x=cal}, h_{x=cal}, \theta_{x=cal})] \varphi_{cam}(\lambda) d\lambda \quad (3.41)$$

If Eq.(3.40) is examined, the linear relation between $DN_{x=cal}$ and the in-band apparent calibration radiance ($L_{app,x=cal}$) is seen since $f(\cdot)$ is a linear invertible one-to-one function. That is,

$$DN_{x=cal} \propto L_{app,x=cal}(T_{x=cal}) \quad (3.42)$$

As DN changes linearly with the in-band apparent radiance, the temperature calibration for all the possible DN values in the linear working regime of the camera can be found. Firstly, $L_{app,x=cal}$ values are computed using Eq.(3.41) for the $DN_{x=cal}$ and $T_{x=cal}$ pairs. Once the $DN_{x=cal}$ and $L_{app,x=cal}$ pairs are obtained, for any DN value in the interval ranging from the minimum $DN_{x=cal}$ value to the maximum $DN_{x=cal}$ value, the respective in-band apparent radiance value, $L_{app,int}$, is found using linear interpolation. Finally, the apparent temperature, T_{int} , for $L_{app,int}$ value is computed using,

$$T_{int} = \arg \min_T |L_{app,x=cal}(T) - L_{app,int}| \quad (3.43)$$

3.2.3.1 Practical Issues

At the end, apparent temperature and DN pairs are obtained that are T_{int} and DN . The pairs are obtained for each possible DN value in the dynamic range of the camera. The search space of the temperature values in Eq.(3.43) are constructed so that the temperature step is less than the NETD value of the camera. Hence, an appropriate temperature accuracy is obtained. Moreover, temperature calibration with radiometric interpolation process is applied only once for an IR camera measurement. The pairs are stored as a look-up table. Then, for each frame in the recording, the look-up table is used to convert the DN values into apparent temperature values with a very low computational complexity. That is temperature values are obtained simply by getting the value in the look-up table using DN for the respective index.

CHAPTER 4

DATA FUSION

The data from multiple sensors arises the question of which data to process. In this thesis dual band IR measurements taken. To be able to detect the moving vehicles in the image image for vehicle classification, one shall decide which frame to segment. In this context, data fusion is used so that best properties of both frames are merged in a single “super frame”.

Data fusion can be classified into 4 classes. These classes are symbol level fusion, object level fusion, feature level fusion, and pixel level (image) fusion [72]. The mentioned classes are given in the order from the highest level of abstraction to the lowest.

Apart from the classification given above, there exist other schemes. According to [73], data fusion can be classified into 3 classes that are centralized data (information) fusion, distributed data (decision) fusion, and feature fusion. According to [74], there are 4 data fusion classes. These are symbol level fusion, feature level fusion, pixel level (image) fusion, and signal level fusion. Saeedi et. al. [75] classifies data fusion into decision fusion, feature fusion, and pixel level image (signal level image) fusion.

In this thesis, the classification of [72] is used. Moreover, it is clear from the given classes above, information fusion is equivalent to symbol level fusion [76] and decision fusion is equivalent to object level fusion. Signal level fusion is not considered as a class of data fusion.

4.1 Symbol Level (Information) Fusion

Symbol level fusion or information fusion uses symbols to describe the information in input images and uses these symbols for data fusion [77]. Symbol level fusion uses the information from multiple sources at the highest level of abstraction [74].

4.2 Object Level (Decision) Fusion

Object level fusion or decision fusion is applied by defining and extracting certain, predefined shapes like human faces or trees [72].

[78] proposes a Principle Component Analysis (PCA) method based face recognition system. Two independent PCA systems are trained at first to get the eigenface images, one with the front views and the other with the depth maps. In the test phase, the front view and the depth map of the face of a test person is fed into two independent systems. Each system returns the Euclidean distances of the test person to the persons in the database. Then the distances returned by each system are multiplied. The person in the database with the smallest distance is the classification result. Since the two systems yield independent classification results and the final decision is given after the fusion of the classification results; this system is an example of a decision fusion system. According to [79], Fisherface methods improve the performance of direct PCA approach by applying PCA for dimensionality reduction and then Fisher's Linear Discriminant Analysis (LDA). LDA extracts features that are more suitable for classification purposes, i.e., most discriminating features, while eigenfaces find the most expressive features, which are not necessarily helpful for recognition.

4.3 Feature Level Fusion

In feature level fusion, the given multiple feature sets are used to produce new fused feature sets, which are more helpful in the final classification.

In [80], Generalized Karhunen Loève Transform (GKLT) is introduced. With the use of GKLT, two sets of different features are fused. The fused feature set is shown to

have a superior performance than the one with uncombined feature sets. [73] is an extension of [80], where the classification with the fused features sets is called parallel strategy and the classification with the uncombined feature sets is called serial strategy. Similarly, it is shown through experiments on some databases, parallel strategy is better than serial strategy.

Feature level fusion is subdivided into two categories, feature selection based and feature extraction based fusion [73].

4.3.1 Feature Selection Based Fusion

In the feature selection based fusion, all feature sets are first grouped together and then a suitable method is used for feature selection [73].

4.3.2 Feature Extraction Based Fusion

In the feature extraction based fusion, the multiple feature sets are combined into one set of feature vectors that are input into a feature extractor for fusion [73].

4.4 Pixel Level (Image) Fusion

Image fusion methods are used to combine multiple pre-registered input images from different sensors or from different exposures into a single fused image [77]. In [72], image fusion is divided into non-multiscale decomposition and multiscale decomposition (transform) based fusion categories. On the other hand, in [77], image fusion is divided into space domain based, transform domain, and multiscale decomposition based fusion categories. Since the latter case is more general, in this thesis the categorization of [77] is used.

4.4.1 Space Domain Based Image Fusion

Space domain based methods perform image fusion in the space domain without using any transform or decomposition operator [77].

Pixel averaging, max-min method, color composite (false color), PCA method, and entropy based fusion are the examples of the space domain based methods.

4.4.1.1 Pixel Averaging Based Image Fusion

Pixel averaging based image fusion is used in [81, 82]. In this method, the fused image is simply obtained by averaging the input images. Fusion operation is done for every pixel of the fused image independently.

$$I_f(x, y) = \frac{\sum_{i=1}^N w(i)I_i(x, y)}{\sum_{i=1}^N w(i)} \quad (4.1)$$

where I_f is the fused image, x and y are the row and column indices of the relevant image respectively, N is the number of the input images, w is the weighting coefficients of the averaging, and I_i is the i^{th} input image [72, 77].

4.4.1.2 Max-Min Method Based Image Fusion

Non-linear maximum or minimum operation is used in the Max-Min based image fusion. The choice of maximum or minimum depends on the application. In the Max-Min method, fused image is obtained by taking the maximum or minimum of the respective images. Fusion operation is done for every pixel of the fused image independently [72, 82].

$$I_f(x, y) = \max [I_1(x, y), I_2(x, y), \dots, I_N(x, y)] \quad (4.2)$$

or

$$I_f(x, y) = \min [I_1(x, y), I_2(x, y), \dots, I_N(x, y)] \quad (4.3)$$

4.4.1.3 Color Composite (False Color) Based Image Fusion

In false color fusion techniques, raw or processed input images are assigned to one of the three color channels (red, green, and blue) of the fused image. Therefore, false color methods are limited to 3 input images maximum [72].

An example of the color composite based fusion algorithm is given in [72]. In this algorithm minimum, maximum, and arithmetic average of the two input images is computed. Minimum image is assigned to red channel, maximum image is assigned to green channel and arithmetic average image is assigned to blue channel of the fused image.

In [81], false color fusion is used. In this study, two input images are fused. First input image is assigned to red channel of the fused image. Second input image is assigned to both green and blue channels of the fused image.

4.4.1.4 PCA Method Based Image Fusion

In the PCA method based fusion, the optimal coefficients in terms of information content and redundancy elimination are calculated using Karhunen - Loéve transform of the intensities. The coefficients for each source image are obtained from the normalized eigenvector associated with the largest eigenvalue of the covariance matrix of the source images [72].

In [81], PCA method based fusion is used. From two input images, two PCA images are obtained. Then, the first PCA image is assigned to red channel of the fused image. First and second input images are assigned to green and blue channels of the fused image, respectively.

4.4.1.5 Entropy Based Image Fusion

In entropy based fusion, input images are divided into blocks. The probabilities of each pixel value is computed by the histogram of the input images. Then, the entropy of each block of each input image is computed. The fused image is obtained by

smoothing the combined blocks that have the greatest entropy values among respective input image blocks [77].

4.4.2 Transform Domain Based Image Fusion

Transform domain based methods perform image fusion in the transform domain [77].

Discrete Cosine Transform (DCT) based image fusion is an example of the transform domain based methods.

4.4.2.1 DCT Based Image Fusion

In DCT based fusion, 2D discrete cosine transforms of sub-blocks of each input image is computed. Then the fused DCT domain image is obtained by fusing the DC and AC coefficients of the cosine transforms of each block. The simplest approach is fusing the DC coefficients by taking the average of the DC coefficients of the related blocks of each input image. AC coefficients are computed by taking the maximum AC coefficient of the related blocks of each input image for every pixel of the block. Finally, the fused image is computed after taking the inverse 2D DCT [72].

In [72], a DCT based fusion algorithm for two input images is proposed. In the fusion process, the DC coefficients are the arithmetic average of the related DC coefficients of the input images. AC coefficients are computed by taking the weighted average of the related AC coefficients of the input images. The weighting coefficients are the L_2 norms of the respective blocks. The use of L_2 norm prevents discontinuities at the edges.

[77] proposes a similar DCT based fusion algorithm. DC coefficients are computed the same way. AC coefficients are computed by taking the weighted average of the related AC coefficients of the input images for each block. The weighting coefficients are unity except the one with the maximum AC coefficient. The coefficient of the block with the maximum is set to 3.

4.4.3 Multiscale Decomposition Domain Based Image Fusion

In multiscale decomposition based methods input images are decomposed into multiscale representations. Then the fusion process is applied and the final fused image is obtained by applying the inverse multiscale decomposition operator [77].

Pyramid transform based and wavelet transform based image fusion are the examples of the multiscale decomposition domain based methods.

4.4.3.1 Pyramid Transform Based Image Fusion

Pyramid transform based fusion techniques are divided into 6 sub-classes. These are Laplacian, Filter-Subtract-Decimate, Ratio (of Low Pass), Contrast, Gradient, and Morphological Pyramid Transform based fusion techniques. In pyramid based fusion, the input images are decomposed and then the fusion process is applied to the decomposed images to get the decomposed fused image. The fused image is obtained by taking the inverse transform of the decomposed fused image. In the fusion process, mostly the mean value is used for the lowest level of the pyramids. For the remaining levels of the pyramids, the maximum value is generally used [72].

In [83], a pyramid transform based fusion approach is proposed. Laplacian pyramids are constructed using the Filter-Subtract-Decimate method. The proposed method is used for extending the depth of field for color images and extending the dynamic range for multiple color images.

4.4.3.2 Wavelet Transform Based Image Fusion

In wavelet transform based fusion, input images are transformed using either shift variant or shift invariant wavelet transforms. For the lowest level Low-Low (LL) component of the fused image in the wavelet domain, the mean value among all the inputs is generally used for the fusion process. For the Low-High (LH), High-Low (HL), and High-High (HH) components of the fused image in the wavelet domain at a given layer, the pixel with the maximum value among all the inputs is selected. Hence the LL components are calculated by the arithmetic average, whereas the other

bands are calculated with the local maximum method. The fused image is obtained by taking the inverse wavelet transform of the wavelet domain fused image [72]. The same approach is used in [81, 84], too.

[74] proposes a wavelet transform based fusion algorithm. Wavelet transforms of the input images are computed. Then the wavelet transform of the fused image is obtained by taking the value with the maximum absolute value at each point. Hence, the salient features such as edges, lines, and region boundaries in the input images are preserved. Moreover, a modification is also proposed. In this approach, majority vote in a local window is used to modify the fusion result if necessary. For instance, if the surrounding pixels of the current pixel in the local window (3x3 or 5x5) come from the first image but the current pixel come from the second image, as the majority vote suggests, the current pixel value is changed with the one of the first image. Hence the fusion algorithm is an area-based maximum selection rule with the consistency verification step.

[85] proposes a wavelet transform based fusion algorithm. Especially, different approaches to merge the wavelet coefficients is given in order to get the fused image. Merging strategies are grouped as activity-level measurement, coefficient grouping method, and coefficient combining method. Moreover, coefficient combining method is sub-divided into selection, general weighted average, adaptive weighted average, fusion by energy comparison, region-based fusion by multiresolution feature detection, background elimination, and variance area based. As an example, fusion by maximum value selection in [74] is a member of the selection sub-group of the coefficient combining method group.

According to [86], the recognition performance depends on the utilized wavelet functions. The performance of compactly supported wavelet is better than that of non-compactly supported wavelet. The performance of symmetrical wavelet is better than that of non-symmetrical wavelet. In addition, using wavelet with proper vanishing moments could improve the recognition rate.

In [87], hyperspectral images are classified using Gabor feature extraction for each pixel. Firstly, Gabor wavelet transforms of the hyperspectral images are computed for different angles, scales, and spectral bands. Genetic algorithm is used to find the

3D transform points. Pruning is applied in the search space. As the final step, the obtained features are used for the classification. Support Vector Machines (SVM) and 3 Nearest Neighbors (NN) classifiers are used.

[75] proposes a dual tree discrete wavelet transform based fusion algorithm. Dual tree discrete wavelet transform of two input images are computed. The low frequency components are fused with a population-based optimization rule that is either particle swarm optimization or genetic algorithm. The high frequency components are fused with a fuzzy-based approach. Fuzzy logic is used to integrate the outputs of three fusion rules, which are weighted averaging, selection using pixel based decision map, and selection using region based decision map. Hence, the advantages of pixel based and region based approaches are utilized in a single scheme.

4.5 Utilized Data Fusion Method

In practice, data fusion is needed only to help the following moving vehicle detection process. Therefore, in applications where computation time is critical, data fusion step may be dropped assuming the vehicle detection is done directly with the MWIR rather than fused frames. However, to simplify the vehicle detection problem and also for the sake of completeness of the system, data fusion is utilized in this thesis. The wavelet transform based image fusion method that is given in [74] is used with Haar wavelets. The related method has a competitive performance and also it has a moderate computational complexity.

Besides, we propose to use a normalization scheme before image fusion. As pointed out in Sec.(3.2), DN values directly depend on the integration time being used at the camera. For instance, assume that an object produces DN_{obj} value at the camera using an IT value, IT_{obj} . If IT_{obj} value is increased, DN_{obj} value increases, too. Therefore, to make the input frames of data fusion independent of the IT values being used on the cameras, the input frames are normalized with the minimum and maximum DN values corresponding to the radiometric calibration limits. Hence, a radiometric bond between the cameras is obtained in terms of calibrated blackbody temperature and apparent radiance values. Although, it has no effect on the sequential processes, the

fused image, which has pixel values in $[0, 1]$, is rescaled with the global maximum and global minimum DN values of the radiometric calibration limits. The last step is just for image visualization purposes.

CHAPTER 5

MOVING VEHICLE DETECTION

To extract the features that will be used for vehicle classification, moving vehicle blobs shall be obtained. For this purpose vehicle detection is used, which takes either the *DN* or fused image as input. The process consists of background subtraction and post-processing blocks and it outputs blobs that are in motion.

5.1 Background Subtraction

The initial step of the moving vehicle detection is the separation of the background and the moving object(s). Basically, the input image is taken and a foreground mask is yielded at the output. A literature review about the topic is given in the following subsections together with the feasibility considerations of each method for our system.

5.1.1 Frame Differencing

In this method, background model is equal to the previous frame. The difference between current frame and previous frame is thresholded and designated as the foreground mask. That is,

$$M(x, y, t) = \begin{cases} 1, & \text{if } |I(x, y, t) - I(x, y, t - 1)| \geq \xi \\ 0, & \text{otherwise} \end{cases} \quad (5.1)$$

where I is the image, x and y are the respective horizontal and vertical image indices, t is the time, and ξ is the threshold. M is the foreground mask [2, 3, 6].

The method has a very low computational cost. However, it cannot cope with multi-modal distributions, illumination changes, IT changes, and periodic movements such as the motion of the leaves of trees [88].

5.1.2 Moving Average Filtering

In this method, foreground mask is obtained by averaging previous N frames. That is,

$$M(x, y, t) = \begin{cases} 1, & \text{if } |I(x, y, t) - I_{bg}(x, y, t)| \geq \xi \\ 0, & \text{otherwise} \end{cases} \quad (5.2)$$

where I_{bg} is the background model and it is updated according to the following equation.

$$I_{bg}(x, y, t) = \alpha I(x, y, t) + (1 - \alpha)I(x, y, t - 1) \quad (5.3)$$

where α is the learning parameter and it determines how the background model adapts to the changes in the scene [4, 7, 10, 11].

Although this method is superior to the frame differencing method, it cannot cope with the multi-modal distributions, too. Moreover, threshold value is very critical on the performance of the algorithm. Another drawback is that the method may produce tails at the back of the moving objects due to the motion of the object in the previous frames [88].

5.1.3 Single Gaussian

In this method, a Gaussian distribution is tried to fit to each pixel. The threshold value is dynamically changed. The foreground mask is obtained with,

$$M(x, y, t) = \begin{cases} 1, & \text{if } |I(x, y, t) - \mu(x, y, t)| \geq 2.5\sigma(x, y, t) \\ 0, & \text{otherwise} \end{cases} \quad (5.4)$$

where μ is the mean of the Gaussian distribution and σ is the standard deviation. Parameters of the distribution is updated as,

$$\mu(x, y, t) = (1 - \alpha)\mu(x, y, t - 1) + \alpha I(x, y, t) \quad (5.5)$$

$$\sigma^2(x, y, t) = (1 - \alpha)\sigma^2(x, y, t - 1) + \alpha[I(x, y, t) - \mu(x, y, t)]^2 \quad (5.6)$$

where α is the learning rate. When α is high, recent pixel values have more influence on the background model [13, 14].

The method is superior to the previous methods. However, it still cannot cope with multi-modal distributions, since single Gaussian is not enough in these cases [88].

5.1.4 Mixture of Gaussians

In [5, 17, 18], Mixture Of Gaussians (MOG) is used to detect the vehicles. The method can deal with periodic clutter, lighting, and IT changes at the cost of increased computational complexity [88].

In [89], the number of Gaussian components for each pixel is the same and constant over time. However, this approach is not optimal in terms of detection and computation time. [90] proposes an online algorithm that estimates the parameters of MOG and simultaneously selects the number of Gaussian components using Dirichlet prior. Therefore, number of Gaussian components is dynamically adapted to the multi-modality of each pixel [88]. Moreover, [90] solves the problem of initial weight calculation [91].

In [90, 92], a Gaussian Mixture Model (GMM) based background subtraction method is proposed. The method aims to automatically adapt to the scene and chooses the right number of components for each pixel based on a model selection criterion from a Bayesian perspective.

In the following equations of the GMM method, the vector notations of the Red Green Blue (RGB) color space is dropped since in this thesis the recordings are gray scaled. Let $x(t)$ be the value of a pixel at time t , then the Background (BG) and Foreground

(FG) decision criterion is given as

$$\frac{p[BG|x(t)]}{p[FG|x(t)]} = \frac{p[x(t)|BG]p(BG)}{p[x(t)|FG]p(FG)} \underset{x(t) \in FG}{\underset{x(t) \in BG}{\geq}} 1 \quad (5.7)$$

In the study, $p[x(t)|FG]$ is assumed to be constant since there is no information about foreground. Then the equation simplifies to

$$p[x(t)|BG] \underset{x(t) \in FG}{\underset{x(t) \in BG}{\geq}} c_{thr} \quad (5.8)$$

where

$$c_{thr} \triangleq \frac{p[x(t)|FG]p(FG)}{p(BG)} \quad (5.9)$$

If Ω denotes the training set for the background model, the estimated BG model is $p[x(t)|\Omega, BG]$. The study assumes no correlation between the adjacent pixels, which means it is a pixel based approach. There are methods, which introduce correlation e.g. Markov random field or filtering. These methods may improve the results but at the cost of a slower performance [90, 92].

The mentioned GMM approach uses an adaptation period, namely T . Hence, the training set becomes $\Omega_T = \{x(t), \dots, x(t - T)\}$. For each new sample, Ω_T is updated and the density is re-estimated, i.e. $\hat{p}[x(t)|\Omega_T, BG + FG]$. The estimated density is assumed to be a mixture of Gaussian densities.

$$\hat{p}(x|\Omega_T, BG + FG) = \sum_{m=1}^M \hat{\pi}_m \mathcal{N}(x; \hat{\mu}_m, \hat{\sigma}_m^2) \quad (5.10)$$

where M is the number of Gaussian components, $\mathcal{N}(\cdot)$ is the Gaussian probability density function, $\hat{\pi}_m$ is the estimated mixing weight, $\hat{\mu}_m$ is the mean estimate, and $\hat{\sigma}_m^2$ is the variance estimate of the m^{th} component. In order to have a legitimate density, the following constraints shall be satisfied [90, 92].

$$\hat{\pi}_m \geq 0 \quad \forall m \quad (5.11)$$

and

$$\sum_{m=1}^M \hat{\pi}_m = 1 \quad (5.12)$$

The iterative update equations for the estimation of the mean and variance values of

each component for every pixel, the following equalities are used.

$$\hat{\pi}_m \leftarrow \hat{\pi}_m + \alpha[o_m(t) - \hat{\pi}_m] \quad (5.13)$$

$$\hat{\mu}_m \leftarrow \hat{\mu}_m + o_m(t) \frac{\alpha}{\hat{\pi}_m} [x(t) - \hat{\mu}_m] \quad (5.14)$$

$$\hat{\sigma}_m^2 \leftarrow \hat{\sigma}_m^2 + o_m(t) \frac{\alpha}{\hat{\pi}_m} \{[x(t) - \hat{\mu}_m]^2 - \hat{\sigma}_m^2\} \quad (5.15)$$

where α is the exponentially decaying envelope and it is suggested to be used as $\alpha = T^{-1}$. $o_m(t)$ is the ownership function of the m^{th} component. $o_m(t)$ is set to 1 for the “close” component with the largest $\hat{\pi}_m$ and others are set to 0. A sample is said to be “close” to a component if the Mahalanobis distance from the m^{th} component is less than a threshold. The squared distance from the m^{th} component is calculated as

$$D_m^2(x(t)) = \frac{[x(t) - \hat{\mu}_m]^2}{\hat{\sigma}_m^2} \quad (5.16)$$

If there are no “close” components, a new component is generated with $\hat{\pi}_{M+1} = \alpha$, $\hat{\mu}_{M+1} = x(t)$, and $\hat{\sigma}_{M+1}^2 = \sigma_0^2$ where σ_0^2 is an appropriate initial variance value. If the maximum number of components is reached, the component with the smallest $\hat{\mu}_m$ value is discarded [90, 92].

The density estimate in Eq.(5.10) contains models for both FG and BG . Usually, the intruding FG objects are represented by some additional clusters with small $\hat{\pi}_m$ values. Hence, the density for the BG can be estimated using the first B largest clusters that is

$$\hat{p}(x|\Omega_T, BG) \approx \sum_{m=1}^B \hat{\pi}_m \mathcal{N}(x; \hat{\mu}_m, \hat{\sigma}_m^2) \quad (5.17)$$

If the components are sorted to have descending $\hat{\pi}_m$,

$$B = \arg \min_b \left[\sum_{m=1}^b \hat{\pi}_m > (1 - c_f) \right] \quad (5.18)$$

where c_f is a measure of the maximum portion of the data that can belong to FG objects without influencing the BG model [90, 92].

The update equation of the mixing weights in Eq.(5.13) is based on a Maximum Likelihood (ML) estimate. In the study, a Maximum A Posteriori (MAP) estimate for large T values is proposed. In this improved estimate,

$$\hat{\pi}_m \leftarrow \hat{\pi}_m + \alpha[o_m(t) - \hat{\pi}_m] - \alpha c_T \quad (5.19)$$

where after each update mixing weights need to be normalized that is $\sum_{m=1}^M \hat{\pi}_m = 1$. Moreover, the component m is discarded if $\hat{\pi}_m < 0$ [90, 92].

In summary, GMM background subtraction method is summarized as,

- Use Eq.(5.17) to classify the new sample as background if $\hat{p}[x(t)|\Omega_T, BG] > c_{thr}$.
- Use Eq.(5.19), Eq.(5.14), and Eq.(5.15) to update $\hat{p}(x|\Omega_T, BG + FG)$.
- Use Eq.(5.18) to select the component of the GMM that belongs to the BG and update $\hat{p}(x|\Omega_T, BG)$.

5.1.5 Non-parametric Approach

In [92], a second background subtraction method that is called non-parametric approach is proposed. The proposed method is an extension of the work [93]. In the method, the authors use balloon variable-size kernel approach. In addition, they use uniform kernels for simplicity.

Density estimation using a uniform kernel starts by counting the number of samples k from the data set Ω_T that lie within the volume V of the kernel. The volume V is a hypersphere with diameter D

$$\hat{p}_{N-P}(x|\Omega_T, BG + FG) = \frac{1}{TV} \sum_{m=t-T}^t \mathcal{K} \left[\frac{|x(m) - x|}{D} \right] = \frac{k}{TV} \quad (5.20)$$

where \mathcal{K} is the kernel and

$$\mathcal{K}(u) = \begin{cases} 1 & \text{if } u < 0.5 \\ 0 & \text{otherwise} \end{cases} \quad (5.21)$$

The volume V of the kernel is proportional to D^d where d is the dimensionality of the data. In practice, kernel \mathcal{K} has little influence but the choice of D is critical [92].

The so called ‘‘balloon estimator’’ adapts the kernel size at each estimation point x . Instead of global optimum D , variable size D for each new point x is used so that a fixed amount of data k is covered. Hence, smaller kernels at densely populated

areas are obtained and vice versa. However, this estimate is not a proper density since it does not sum up to one. Still, balloon estimate is often used for classification problems as it is related to k-NN classification. In the study $k = [0.1T]$, where $[.]$ is round-to-integer operator, is used. $c_{thr} \propto V^{-1} \propto D^{-1}$ since in the balloon variable kernel approach k is fixed and V is variable [92].

In practice, T is large and keeping all samples in Ω_T requires too much memory. Hence, a fixed number of samples $K \ll T$ is used. The model is split into “short-term” (K_{short} samples from T_{short} period) and “long-term” model (K_{long} samples from T_{long} period). Without proof of optimality, $K_{short} = K_{long} = 0.5K$ is used in the study. The authors choose the “short-term” model to approximate the first 30% of the information under the envelope, i.e. $T_{short} = \frac{\log 0.7}{\log(1-\alpha)}$.

$$\hat{p}_{N-P}(x|\Omega_T, BG) = \frac{1}{TV} \sum_{m=t-T}^t b(m) \mathcal{K} \left[\frac{|x(m) - x|}{D} \right] \quad (5.22)$$

where indicator $b(.)$ is equal to 0 if the sample is assigned to FG and equal to 1 if it is assigned to BG . Hence, the BG model considers only the samples assigned to BG . If $\hat{p}_{N-P}(x|\Omega_T, BG) > c_{thr}$, pixel is classified as BG . Eq.(5.20), in which all samples (regardless of $b(.)$'s) are considered, is used to determine $b(.)$ for the new sample [92].

In summary, non-parametric background subtraction method is summarized as,

- Use Eq.(5.22) to classify the new sample $x(t)$ as background if the condition $\hat{p}[x(t)|\Omega_T, BG] > c_{thr}$ holds.
- Add $x(t)$ to Ω_T and remove the oldest sample to update $\hat{p}(x|\Omega_T, BG + FG)$.
- If Eq.(5.20) $> c_{thr}$ set $b(m) = 1$ for the sample, i.e. use the new sample for $\hat{p}(x|\Omega_T, BG)$, to update $\hat{p}(x|\Omega_T, BG)$.

If K is increased, better segmentation is obtained but at the cost of increased processing time [92].

5.1.6 GraphCut

GraphCut combines the two already known approaches for image segmentation algorithms based on colors (or more precisely gray-levels) and the contrast in different regions of an image. For successful segmentation, the energy formulation

$$E(z) = P(z) + \gamma C(z) \quad (5.23)$$

has to be minimized. The weighting parameter γ controls the importance of one term over the other. The fidelity term $P(\cdot)$ gives rise to a cost function, which penalizes false classification of a pixel z to the foreground ($\alpha = 1$) or to the background ($\alpha = 0$). Since the user provides a so-called trimap, where two regions “sure foreground” and “sure background” has to be defined, one can easily calculate a probability distribution and cost functions $p_{z,\alpha}$ from the gray-valued pixels and the image histograms of these two regions

$$P(z) = \sum_{z \in I} p_{z,\alpha} \quad (5.24)$$

where I is the image. Costs can be calculated from the negative log-likelihood of the probability belonging either to the foreground or to the background.

A prior term $C(\cdot)$ representing the pairwise interactions between neighboring pixels is calculated from the contrast between each two neighboring pixels z and \hat{z} with

$$C(z) = \sum_{(z,\hat{z}) \in N} c_{z,\hat{z}} \quad (5.25)$$

where the neighborhood N is chosen such that only neighboring pixels around the segmentation boundary are summed up. These are the only pixels, z and \hat{z} , belonging to two different foreground/background maps: $\alpha_z \neq \alpha_{\hat{z}}$. Only a 4-way neighborhood is used in the study. Therefore, the minimization criterion is to find the shortest possible segmentation border that gives the smallest sum over its contrast terms. The contrast between neighboring pixels z and \hat{z} can be expressed as

$$c_{z,\hat{z}} = \exp \left[-\frac{(I_z - I_{\hat{z}})^2}{2\sigma^2} \right] \quad (5.26)$$

where I_z is the gray-value of the pixel z in the range $0, \dots, 1$. The variance σ^2 over all differences in intensity can be seen as the noise floor present in the image. Choosing

this parameter carefully lets the contrast term successfully switch between almost zero for high contrast and one vice versa. However, other functions, separating noise from real contrast in the same manner, are also possible. From these two properties of each pixel - one belonging to the object or the background, the other being an edge or not - an undirected graph is built. More precisely a so called S/T - *graph* is built, where the two terminals S and T represent the object and the background, respectively. Edges from and to these terminals are weighted with the corresponding foreground/background costs $p_{z,\alpha}$. Neighboring pixels are connected with edges in 4-way neighborhood, weighted with the corresponding contrast terms $c_{z,\hat{z}}$. Finally using a standard Minimum-Cut/Maximum-Flow (MCMF) algorithm has been proven to give the optimal segmentation border in terms of the energy formulation $E(\cdot)$ defined in Eq.(5.23). The segmentation border corresponds to the edges representing the minimum cut in the graph [94].

Although GraphCut has competitive performance, its main drawback is the need of a manual “sure background” and “sure foreground” selection.

5.1.7 GrabCut

GrabCut extends the GraphCut to color images. Instead of gray-level histograms, it makes use of GMM. Background and foreground are each described with five full-covariance Gaussian components $M_{z,k}$. So the fidelity term $P(\cdot)$ is now calculated from the superposition of the Gaussian components

$$M_{z,k} = \frac{1}{2\pi\sqrt{\Sigma_k}} \exp \left[-\frac{1}{2}(I_z - \mu_k)^T \Sigma_k^{-1} (I_z - \mu_k) \right] \quad (5.27)$$

where the term I_z reflects a three-valued RGB color of the pixel z . μ_k are the mean color of each component and Σ_k are full-covariance matrices reflecting color dependencies between the three color layers. Adaptation of the probability distributions $M_{z,k}$ to the RGB colors is carried out with the iterative Expectation Maximization (EM) algorithm, according to a predefined trimap given by the user. Due to the 3D color space, the contrast $c_{z,\hat{z}}$ is calculated as

$$c_{z,\hat{z}} = \exp \left(-\frac{\|I_z - I_{\hat{z}}\|^2}{2\sigma^2 \|z - \hat{z}\|} \right) \quad (5.28)$$

where the norm $\|I_z - I_{\hat{z}}\|$ is the Euclidean distance in RGB space and $\|z - \hat{z}\|$ indicates the spatial (Euclidean) distance between two neighboring pixels z and \hat{z} . GrabCut uses a 8-way connectivity. In this manner, the whole algorithm is laid out in an iterative way; after each EM iteration, an S/T - *graph* is built up like in the GraphCut and solved with the minimum-cut algorithm. The resulting segmentation border is used to update the trimap describing foreground and background regions. This new trimap is used for the next EM iteration and so on. The alternating usage of EM steps and MCMF solutions guarantees the proper monotonic energy minimization over time. The amount of changes in the overall energy $E(\cdot)$ between two iterations is used as the stopping criterion for the algorithm [94].

Similar to the GraphCut, the main disadvantage of GrabCut is again the need for the user to select “sure background” and “sure foreground” regions.

5.1.8 GrayCut

[94] aims to segment the gray-valued IR images of a ship. Contrast and color information are combined into an energy minimization criterion. The advantages of GraphCut, which is gray-level-based, and GrabCut, which uses an iterative optimization scheme, algorithms are combined and named as GrayCut algorithm. As in GrabCut, Gaussian mixture models are used, but only to find the distributions in the two gray-scale histograms - the one for the user-defined background and the one for the (unknown) rest. The possible range of values is reduced from the 3D space of RGB colors to the purely one dimensional gray-scale histogram. Adaptation of the Gaussian mixture models is carried out by EM, so the whole algorithm is of an iterative nature. Starting with a random distribution for EM learning as in GrabCut is reported to be not a good starting point for the segmentation task. Hence, the very first EM step is applied before the whole algorithm starts. This guarantees a proper initial distribution of the mixture models, but also ensures the adaptation to changes in the trimap based on intermediate segmentation results. Since the possible range of values and the total number of components has been reduced, the overall algorithm performance has been slightly increased. Moreover, less iterations are needed for the Gaussian components to adapt the gray-level histogram. However, an important and

critical manual step used in the algorithm is the requirement of the user to select the background area.

5.1.9 Object-based Segmentation

Object-based segmentation methods are generally based on using 3D models of the objects. The detection is done by comparing 3D models of the objects with the input frames. Clearly, these methods have higher computational complexity [88]. [1, 8] use object-based segmentation to obtain the vehicle blobs.

5.2 Utilized Background Subtraction Method

In this thesis, GMM background subtraction method proposed in [92] is used. The motivation is based on the results of the related work that GMM performs better than the non-parametric approach for the traffic sequence. Moreover, GMM does not need the background area selection by the user unlike GrayCut, GraphCut, and GrayCut. GMM also needs a fair computation time while having a good performance, when compared to the basic background subtraction methods such as moving average filtering and also to the object-based segmentation methods.

5.3 Post-processing After Background Subtraction

The raw output of the background subtraction process may be noisy. In order to remove the noise present in the foreground mask and also to prepare the data for classification, morphological operations are applied.

In [91], first dilation then erosion, that is closing, operators are applied to the raw foreground mask data. Then, median filtering is utilized. Finally, connected component labeling is used.

In [88], first erosion then dilation, that is opening, operators are applied to the raw foreground mask data. Connected component labeling is used to label the blobs. As

the last step, area filtering is utilized to remove the residual noise. The area of each label is thresholded with the predetermined average area of the scene.

5.3.1 Erosion

Erosion operator is utilized to get rid of the region boundaries of the foreground pixels. A structuring element such as the one given in the following equation can be used.

$$SE_{eros} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (5.29)$$

At the end, the foreground pixels, which are not completely surrounded by other foreground pixels, are relabeled as background pixels. Hence, small regions corresponding to noise are eliminated [95].

5.3.2 Dilation

Dilation operator is utilized to enlarge the region boundaries of the foreground pixels. A structuring element such as the one given in the following equation can be used.

$$SE_{dilat} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (5.30)$$

At the end, the background pixels, which are not completely surrounded by other background pixels, are relabeled as foreground pixels. Hence, small holes due to the background subtraction artifacts are filled back [95].

5.3.3 Closing

Closing operator is the combination of dilation followed by erosion [95]. It is used to fill the smaller holes than the structuring element in an object.

5.3.4 Opening

Opening operator is the combination of erosion followed by dilation [95]. It is used to remove noise without affecting the object boundaries.

5.3.5 Connected Component Labeling

Connected component labeling assigns a unique label to each component after grouping the pixels in an image into components based on pixel connectivity. The algorithm is as follows;

- Image is raster scanned.
- For each of the image pixels, if the current pixel is a foreground pixel
 - If one of the pixels on the left, on the top, on the upper left or on the upper right is already labeled, the respective label is assigned to the current pixel, too. This scheme is called 8-connectivity.
 - If two or more of the neighbors have a label, one of the labels is assigned to the current pixel. All of the labels are marked as equal by forming an equivalence table.
 - If none of the neighbors has a label, a new label is assigned to the current pixel.
- Once the raster scan is completed, labels corresponding to the same group of pixels in the equivalence table are merged and replaced by a single label.
- Image is scanned once more to update the labels [96].

5.4 Utilized Post-processing After Background Subtraction Method

In this thesis, in the post-processing first opening followed by closing operators are applied to the FG mask. Then, using connected component labeling with 8 connect-
edness, unique labels are assigned to each component in the segmented image. Area
filtering, with a threshold appropriate for the scene of interest, is used to eliminate
the connected components, which are actually residual noise. Finally, for each of the
remaining connected component, the inner region of the component is filled. That is
achieved by updating the pixels, which lie within the outermost contour of the com-
ponent, as foreground.

As an example, each step of the moving vehicle detection method is analyzed using
a sample *DN* image captured by the LWIR camera of the dual band measurement
setup. The input image and also the segmented image are shown in Fig.5.1(a) and
Fig.5.1(b), respectively.

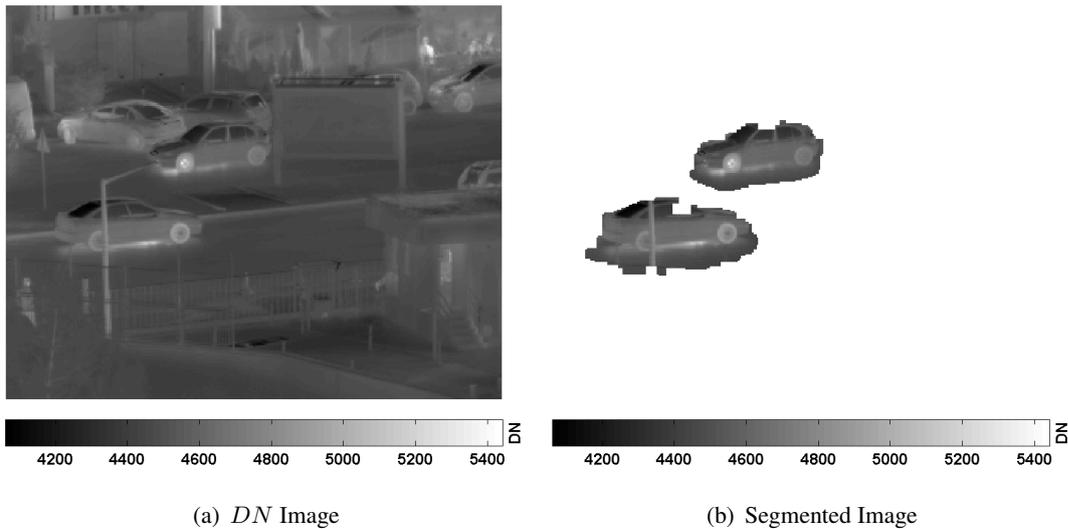
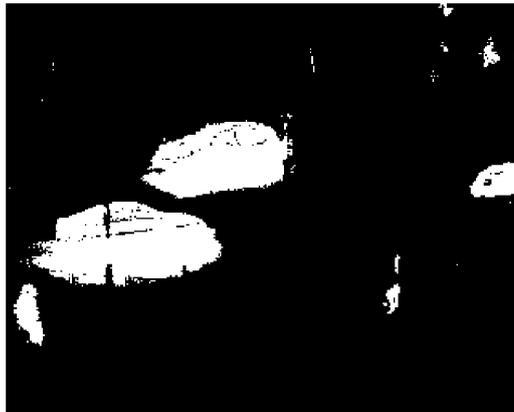


Figure 5.1: A Sample Result for the Moving Vehicle Detection Method

The foreground mask obtained by the utilized BG subtraction method is given in
Fig.5.2(a). The opened and closed image masks are shown in Fig.5.2(b) and Fig.5.2(c),
respectively. The output of the connected component labeling is given in Fig.5.2(d),
where any color other than white represents a connected component. The image mask
obtained after area thresholding is shown in Fig.5.2(e). Finally, the last step of the al-
gorithm that is the filled image mask is given in Fig.5.2(f). Therefore, it is shown that

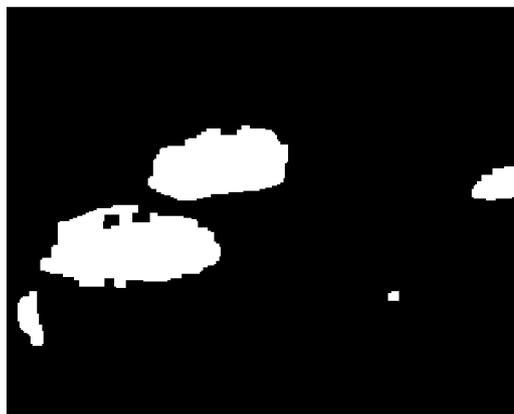
the detection method can successfully separate BG from the moving vehicles with the help of post-processing steps.



(a) Foreground Image Mask



(b) Opened Image Mask



(c) Closed Image Mask



(d) Connected Component Label Image



(e) Area Thresholded Image Mask



(f) Filled Image Mask

Figure 5.2: A Sample Result for the Utilized Post-processing After Background Subtraction Method

CHAPTER 6

CLASSIFICATION

Final step of the proposed system is classification. Classification is defined as the assignment of a new instance to a class, which is a group of previously known instances. Classification process mainly consists of feature extraction and classifier. In feature extraction, information from the new instance, which is called feature, is obtained. Then, a set of features are fed into a machine learning algorithm, which is the classifier [88].

In [2] a metric, which is the ratio of the perimeter to the area of the vehicle blob, is used for classification into the classes human, vehicle or background clutter. The accuracy of the classification is 85%.

[6] uses length, [3,4,17] use both length and height of the vehicles for classification. Height and length are obtained by using the 2D projections of the vehicles, where information about camera parameters and camera location are needed. In [3], the classes are truck and other vehicles. In [4], cars and non-cars classes are used. In [6,17], the classes are long vehicles and short vehicles. The classification accuracy of [3] and [4] are 90% and 70%, respectively.

Size of the bounding box of the blobs and also velocity features are used in [5,9]. The classes of [5] are vehicle and pedestrian whereas the classes of [9] are motorcycle/bicycle, car, and bus/minibus.

In [7] pickup, sedan, van, van truck, truck, trailer, and bus classes are used. Area, size, and length features are used in the classification in which 91% accuracy is obtained.

Size and linearity features, a measure for the roughness of the vehicle silhouette, is used in [12]. Car, minivan, van-truck/bus, and truck classes are used and 88% accuracy is reported.

Blob features are used in [13, 14] such as breadth, area, compactness, perimeter, elongation, roughness, length, long and short axis of the fitted ellipse, centroid, and image moments. LDA is used for dimensionality reduction. The classification accuracy of [13] is 82.9% with the classes sedan, semi, and truck/SUV/van. The accuracy of [14] is 74.4% when the sedan, truck, SUV, semi, van, truck/SUV/van, and moving truck classes are used.

[10, 11, 19–21] use vehicle images for classification. [10, 19–21] use PCA and [11] uses Independent Component Analysis (ICA) for dimensionality reduction of the feature space. Vehicle and non-vehicle classes are used in [19], [20], and [21] with 94%, 95%, and 93.04% classification rate, respectively. Passenger, car, van, and pick-up classes are used in [10] and [11] with 50% and 75% respective classification precision.

Histogram of Oriented Gradients (HOG) features are used in [15, 97]. Car, van, pickup-truck, and no vehicle classes are used in [15], where vehicle and non-vehicle classes are used in [97]. Respective classification accuracies of [15] and [97] are 64.3% and 90%.

In [98], combination of gray-level features and continuity based image descriptors are used. Private cars, mid-sized vehicles, and trucks classes are used and an average correct detection score of 72.35% is achieved.

In [16], bounding boxes of the vehicles are obtained using edges. Then, bounding boxes are verified with corner detection and symmetry inside the box. Vehicle and non-vehicle classes are used in the study and 90% classification rate is reported.

[1, 8, 18] use 3D models of the vehicles for classification. Car and van classes are used in [1] and 96% accuracy is obtained. Bicycle, lorry, motorcycle, van, car, extra-urban bus, urban bus and unknown classes are used in [8] and 91.5% classification rate is achieved. In [18], bus/lorry, van, car/taxi, and motorcycle/bicycle classes are used and 90.4% recall at a precision of 87.9% is reported.

In [22] DN values corresponding to the pre-segmented blobs are compared with the image blocks in the codebook. An Automatic Target Recognition (ATR) is proposed, which consists of cascaded Vector Quantizers (VQ) and Multilayer Perceptrons (MLP). Ten different vehicles classes are used. 90.3% and 65.3% recognition rates are reported for the respective moderately and highly cluttered scenarios.

[24, 25] use PCA to extract features from the DN images of both MWIR and LWIR bands. Then MLP is utilized for target detection. It is concluded that combining MWIR and LWIR band images before the eigenspace transformation yields better detection results for the joint use of dual band data.

In [26], several ATR techniques are compared. These are Convolutional Neural Networks (CNN), PCA, LDA, Learning Vector Quantization (LVQ), Modular Neural Network (MNN), The Hausdorff metric-based matching technique, and a geometric hashing technique. Ten different vehicles classes are used. The respective recognition rates of the techniques are 66.1%, 52.17%, 50.32%, 75.12%, 75.58%, 62.86%, and 50.09%.

In [27] tank, armored personnel carrier, and truck classes are used. Vehicle blobs are found by thresholding on the histogram of the image and area filtering at various ranges. Difference between maximum DN and median DN of the blob, blob size, blob size difference after histogram modification, ratio of the vertical extent of the blob to the horizontal extent, number of new blobs within the adjacency of the blob after histogram modification, and the moving average of the median DN of the blob features are used. 98% Probability of Detection (PD) and 2% Probability of False Alarm (PFA) is reported. Classification is done by a fuzzy-inference system and the classifier is “tuned by hand”, which means the training phase is fully manual.

In [28] a combination of a target detector that is based on the features extracted from the DN image and an eigen-neural based clutter rejecter. Target detector correctly locates 90% of all legitimate targets. 87.54% hit rate is reported for the clutter rejecter.

[29] uses PCA to perform dimensionality reduction and also a supervised MLP for classification. Data fusion of dual band frames are done either at pixel level or at feature level. Dual band DN frames correspond to 12 bit MWIR and LWIR frames.

The authors conclude that both fusion schemes give similar results and so the dimensionality reduction on dual band images can be performed independently.

In [99] PCA, subspace LDA, and ICA are compared for dimensionality reduction of the DN frames. Moreover, nearest neighbor and LVQ classifiers are evaluated. Truck, bus, van, minibus, and tanker classes are used. It is reported that ICA performs the best among the feature extraction methods. Furthermore, among the classifiers LVQ is reported to be the superior one based on the recognition rate results, which is 69.3% for LVQ.

[32] detects moving targets using temporal variance analysis. Then, combined tracking and classification is utilized using appearance model based particle filter. Four different types of vehicles are the classes and a total accuracy of 89.07% is obtained.

In [33], vehicle blobs are used to extract HOG features. Orientations are computed in a scale-space framework. Then, the features are matched with the learned templates in a database using Multinomial Pattern Matching (MPM). Finally, to increase the performance further, MPM scores from multiple frames are fused with Sequential Probability Ratio Test (SPRT). In the classifier, vehicle and non-vehicle classes are used. It is reported that while using a single frame for the classification, 52% PD with 0% PFA and 100% PD with 26% PFA is achieved.

In [36], anomaly detection with kernel wavelet Reed-Xiaoli (RX) method is proposed. 12 bit LWIR DN frames are used in the study. The wavelet transform of the image is computed in a number of equal sub-bands. Then, all the sub-bands are arranged into a sub-band image cube. In the image cube, the first band is the base-band where the remaining bands are the sorted higher frequency sub-bands, where sorting is done based on the energy content. Afterwards, the bands with low energy content are discarded. Finally, kernel RX algorithm that is used for hyperspectral anomaly detection is utilized using the dual-window approach.

[100] uses car, van, and truck classes. Firstly, vehicles are detected using a background subtraction algorithm based on MOG. Then, morphological operations are utilized to cope with the artifacts of the detection. The morphological operations are opening, closing, 8 connected component labeling, and thresholding with the pre-

determined minimum number of pixels of a vehicle in the ROI. Afterwards, blob features that are dimension-based descriptors and HOG features that are shape-based descriptors are extracted. Blob features are area, width, and height of the bounding box, ratio of the blob area to the bounding box area (extent), major and minor axis lengths, and the ratio of foreground pixel number to background pixel number in the bounding box. Then, a hierarchical supervised classifier, SVM, is used. A total accuracy of 80.6% and 96.4% total accuracy is achieved for two different scenarios. [88] is an extended study of [100]. In this study, both a hierarchical and a multi-class supervised classifier, SVM, are used. A total accuracy of 91.6% and 98.1% is reported for two different scenarios.

Aside from the given vehicle classification literature, there are studies that evaluates various classifier fusion schemes such as [23, 30, 31, 34]. However, classifier fusion is beyond the scope of this thesis.

6.1 Feature Extraction

Once the vehicle blobs are obtained, the following process is feature extraction that ideally produces similar values for the instances of a class. In computer vision, there are key point (feature) detectors like Harris [101], Hessian [102], Laplacian Of Gaussians (LOG) [103], Difference Of Gaussians (DOG) [104], Canny [105], Sobel [106], Prewitt [107], Shi and Tomasi [108], Features from Accelerated Segment Test (FAST) [109], Maximally Stable Extremal Regions (MSER) [110], Haar wavelets [111], Hough transform [112–115], and fast radial symmetry transform [116, 117].

Once the features are detected, a local image patch around the feature can be extracted using feature descriptors like MPEG-7 [118], Scale Invariant Feature Transformation (SIFT) [119], Speeded Up Robust Features (SURF) [120], HOG [121], Oriented FAST and Rotated BRIEF (ORB) [122], Gradient Location and Orientation Histogram (GLOH) [123], covariance matrix descriptors [124], and Gabor filters [125]. Among these descriptors, SIFT is a popular one. Its parameters can be modified for individual tasks. This modification is done by the HOG descriptor so

that HOG is more suitable for generic object recognition [98, 124]. HOG algorithm is relatively invariant to local geometric and photometric transformations. This is because rotation and translation do not affect HOG values and illumination invariance is achieved through normalization [88]. Generally as the number of pixels in a cell decreases, the performance tends to increase. However, the size of the HOG feature vector and so the computation time increases, too. Hence there is a trade-off between the computation time and the performance [126]. Moreover, when the HOG feature vector size increases, the computation time of the classifier generally increases, as well. For this reason HOG features are beyond the scope of this thesis.

The classifiers directly depend on the extracted features. Hence, for both training and testing of the classifiers (classification of vehicles), feature vectors are used. In the proposed algorithm, two types of features are used, which are blob features and radiometric features.

6.1.1 Blob Features

One of the discriminative features among vehicle classes are their sizes and appearances. Therefore, for each of the vehicle blob that has a connected component label, the following blob features are extracted.

- Area of the blob
- Ratio of the area of the blob to the area of the bounding box of the blob
- Major axis length, which is a scalar specifying the length (in pixels) of the major axis of the ellipse that has the same normalized second central moment as the blob
- Minor axis length, which is a scalar specifying the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moment as the blob
- Ratio of the number of background pixels to the foreground pixels in the bounding box

- Width of the bounding box
- Height of the bounding box

6.1.2 Radiometric Features

As the source radiance estimate of the ROI is given as input to the classification process (See Fig.(1.2) and Fig.(1.4)), the radiometric features are extracted. The radiometric features are,

- Mean value of the source radiance estimate frame in the ROI computed using

$$\mu_{rad} = \frac{1}{N} \sum_{(i,j) \in ROI} L_{src}(i, j) \quad (6.1)$$

where μ_{rad} is the mean value, N is the number of pixels in the ROI, and i and j are the respective row and column frame indices.

- Standard deviation value of the source radiance estimate frame in the ROI computed using

$$\sigma_{rad} = \sqrt{\frac{1}{N-1} \sum_{(i,j) \in ROI} [L_{src}(i, j) - \mu_{rad}]^2} \quad (6.2)$$

where σ_{rad} is the standard deviation value.

- Skewness value of the source radiance estimate frame in the ROI computed using

$$\varsigma_{rad} = \frac{\frac{1}{N} \sum_{(i,j) \in ROI} [L_{src}(i, j) - \mu_{rad}]^3}{\left\{ \sqrt{\frac{1}{N} \sum_{(i,j) \in ROI} [L_{src}(i, j) - \mu_{rad}]^2} \right\}^3} \quad (6.3)$$

where ς_{rad} is the skewness value.

- Kurtosis value of the source radiance estimate frame in the ROI computed using

$$\kappa_{rad} = \frac{\frac{1}{N} \sum_{(i,j) \in ROI} [L_{src}(i, j) - \mu_{rad}]^4}{\left\{ \frac{1}{N} \sum_{(i,j) \in ROI} [L_{src}(i, j) - \mu_{rad}]^2 \right\}^2} \quad (6.4)$$

where κ_{rad} is the kurtosis value.

- Entropy value of the source radiance estimate frame in the ROI computed using

$$H_{rad} = - \sum_{k=1}^K h_{L_{src}}(k) \log_2 h_{L_{src}}(k) \quad (6.5)$$

where H_{rad} is the entropy value, $h_{L_{src}}$ is the normalized histogram of the ROI of the source radiance frame and K is the number of bins in the histogram.

- Maximum value of the source radiance estimate frame in the ROI
- Minimum value of the source radiance estimate frame in the ROI

6.2 Classifiers

Classifiers assign a new instance to a class by using the extracted feature vector. The assignment relies on training data in supervised learning methods. Machine learning algorithms output classifiers by using the training data, which contains extracted feature vector and the class labels of a known instance set [88].

6.2.1 Nearest Neighbor Classifier

In nearest neighbor classifier, the distance between each vector of the training set and the new feature vector is calculated. The class label of the training vector corresponding to the smallest distance is assigned to the new instance. Although any distance measure can be used, generally Euclidean distance is used that is

$$d(\bar{x}, \bar{y}) = \|\bar{x} - \bar{y}\|_2 = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \quad (6.6)$$

where \bar{x} and \bar{y} are the feature vectors in \mathfrak{R}^m [88].

6.2.2 k-nearest Neighbor Classifier

k-nearest neighbor classifier is an extension to the nearest neighbor classifier. It classifies a feature vector by assigning it to the label that is most frequently used among the k nearest samples [88].

Both nearest and k-NN classifiers do not scale well for large training sets. When the training size increases, the size of the distance vector to be computed for a new instance increases as well. Therefore, computation time increases with the training size [88].

6.2.3 Support Vector Machines

Support vector machines are machine learning method for classification, regression, and other learning tasks. A typical use of SVM for classification consists of two steps, namely training and testing [127]. In the training phase, a data set of the known instances is used to obtain the model. Then, the testing data set is given to the model in order to predict the instances. To achieve the best performance of SVM, the parameter selection is a critical step.

Given training vectors $\bar{x}_i \in R^n$ for $i = 1, 2, \dots, l$ in two classes and an indicator vector $y_i \in R^l$ such that $y_i \in \{-1, 1\}$, the following optimization is solved

$$\begin{aligned} \min_{\bar{w}, b, \xi} \quad & \frac{1}{2} \bar{w}^T \bar{w} + C \sum_{i=1}^l \xi_i \\ \text{subject to} \quad & y_i [\bar{w}^T \phi(\bar{x}_i) + b] \geq 1 - \xi_i, \\ & \xi_i \geq 0, \quad i = 1, 2, \dots, l \end{aligned} \quad (6.7)$$

where l is the size of the testing data set, $\phi(\bar{x}_i)$ maps \bar{x}_i into a higher-dimensional space and non-negative C is the regularization parameter. Since the vector \bar{w} may have high dimensionality, usually the following dual problem is solved

$$\begin{aligned} \min_{\bar{\alpha}} \quad & \frac{1}{2} \bar{\alpha}^T \mathbf{Q} \bar{\alpha} - \bar{e}^T \bar{\alpha} \\ \text{subject to} \quad & \bar{y}^T \bar{\alpha} = 0, \\ & 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, l \end{aligned} \quad (6.8)$$

where $\bar{e} = [1, \dots, 1]^T$ is the vector of all ones, \mathbf{Q} is an $l \times l$ positive semi-definite matrix, $Q_{ij} \triangleq y_i y_j \mathcal{K}(\bar{x}_i, \bar{x}_j)$, and $\mathcal{K}(\bar{x}_i, \bar{x}_j) \triangleq \phi(\bar{x}_i^T \bar{x}_j)$ is the kernel function [127–129].

Once Eq.(6.8) is solved, the optimal support vector \bar{w} satisfies the following relation

$$\bar{w} = \sum_{i=1}^l y_i \alpha_i \phi(\bar{x}_i) \quad (6.9)$$

and the decision function is

$$\text{sgn} [\bar{w}^T \phi(\bar{x}) + b] = \text{sgn} \left[\sum_{i=1}^l y_i \alpha_i \mathcal{K}(\bar{x}_i, \bar{x}) + b \right] \quad (6.10)$$

where $y_i \alpha_i \forall i$, b , the label names, support vectors, and kernel function parameters obtained after training are stored in the model and used for prediction [127].

6.2.3.1 Multi-class SVM

A direct extension of the standard two-class SVM to multi-class classification is the “one-against-all” approach. In this approach, the problem is considered as a set of k two-class problems, where k is the total number of classes. The decision is the argument of the optimal discriminant function with the maximum value. The discriminant function is $(\bar{w}^i)^T \phi(\bar{x}) + b$ where \bar{w}^i is the support vector of the i^{th} class. This technique may lead to indeterminate regions, where more than one discriminant function is positive, which contradicts with the design constraint of the “one-against-all” approach. Moreover, each binary classifier deals with asymmetric problem such that training is carried out with many more negative values than positive examples. This is more serious when k is relatively large [129].

Other methods are available for the multi-class SVM classification other than the “one-against-all” approach. Among these approaches, “one-against-one” approach is adopted in [127] due to its competitive performance. In this approach, $k(k-1)/2$ classifiers are constructed for a total of k classes and each classifier trains data from two classes. For the training of the i^{th} and j^{th} classes, the following two-class classification is solved.

$$\begin{aligned} \min_{\bar{w}^{i,j}, b^{i,j}, \xi^{i,j}} \quad & \frac{1}{2} (\bar{w}^{i,j})^T \bar{w}^{i,j} + C \sum_t (\xi^{i,j})_t \quad (6.11) \\ \text{subject to} \quad & (\bar{w}^{i,j})^T \phi(\bar{x}_t) + b^{i,j} \geq 1 - (\xi^{i,j})_t, \text{ if } \bar{x}_t \text{ is in the } i^{\text{th}} \text{ class} \\ & (\bar{w}^{i,j})^T \phi(\bar{x}_t) + b^{i,j} \leq -1 + (\xi^{i,j})_t, \text{ if } \bar{x}_t \text{ is in the } j^{\text{th}} \text{ class} \\ & (\xi^{i,j})_t \geq 0 \end{aligned}$$

In the classification, voting mechanism is utilized. Each binary classification result is considered as a vote where votes can be cast for all data points. The final decision is given on the basis of majority vote. In the case that two classes have exactly the same number of votes, the first class is designated as the final decision for the sake of simplicity [127, 129].

Standard SVM training algorithms have $O(m^3)$ computational time and $O(m^2)$ computational space complexity, where m is the number of training data [124, 129]. Therefore, with a large size of training data Radial Basis Function (RBF) kernel is preferred due to its lower computational cost with respect to the linear and polynomial kernels. Moreover, with the linear and polynomial kernels, the regularization parameter C can not be adjusted optimally due to the high computational cost [124]. Therefore, in this thesis RBF kernel is used, where RBF kernel takes the following form

$$\mathcal{K}(\bar{x}_i, \bar{x}) = \exp\left(-\frac{\|\bar{x}_i, \bar{x}\|_2^2}{\gamma^2}\right) \quad (6.12)$$

where γ is the kernel parameter of the Gaussian RBF [130].

6.3 Utilized Classification Method

In this thesis, a multi-class SVM classifier with Gaussian RBF kernel is used. SVM is a state of the art classifier and frequently used in vehicle recognition applications. For the reference system only blob features, which is a vector with 7 elements, are extracted and used in the classifier. For the proposed system both blob and radiometric features are extracted and used in the classifier. Hence, in the case of single band system, the feature vector contains 14 elements, whereas it contains 21 elements in the case of dual band system.

6.4 Performance Metrics

The metrics used to evaluate the performance of the proposed classification system are confusion matrix, precision, recall, and accuracy.

6.4.1 Confusion Matrix

Confusion matrices are used to show the number of instances of the predicted and actual classes in the respective columns and rows of the matrix [131]. An example of a confusion matrix for a system with two classes is given in Table 6.1.

Table 6.1: Confusion Matrix Example

		Prediction	
		Class-1	Class-2
Real	Class-1	10	3
	Class-2	5	7

6.4.2 Precision

Precision for class i , in a system with a total k number of classes, is computed using

$$Precision(i) = \frac{\mathcal{A}(i, i)}{\sum_{j=1}^k \mathcal{A}(j, i)} \quad (6.13)$$

where \mathcal{A} is the confusion matrix and precision of class i is the percentage of data points predicted as class i , whose real class label is indeed i [131].

6.4.3 Recall

Recall for class i , in a system with a total k number of classes, is computed using

$$Recall(i) = \frac{\mathcal{A}(i, i)}{\sum_{j=1}^k \mathcal{A}(i, j)} \quad (6.14)$$

where \mathcal{A} is the confusion matrix and recall of class i is the percentage of data points with real class label i , which were correctly predicted in that class [131].

6.4.4 Accuracy

Accuracy is computed using

$$Accuracy = \frac{\sum_{i=1}^k \mathcal{A}(i, i)}{\sum_{i=1}^k \sum_{j=1}^k \mathcal{A}(i, j)} \quad (6.15)$$

where \mathcal{A} is the confusion matrix and accuracy is the percentage of data that has been correctly predicted [131]. In other words, accuracy is a measure of the overall correct decision rate of the classifier.

CHAPTER 7

MEASUREMENT SETUP

Two setups are used in the thesis, one for the single band and one for the dual band system.

7.1 Single Band Measurement Setup

A micro bolometer (See App.(B.3)) LWIR camera is used for the single band case. The spectral range of the camera is from $7.5\mu m$ to $13\mu m$. No filter is used in the camera. The horizontal and vertical frame size is 320 by 240, respectively. Moreover, a $122mm$ lens is used so that horizontal and vertical FOV value is 7 by 5.25 degrees. The camera has 16 bit resolution in the DN values.

7.2 Dual Band Measurement Setup

A cooled photo voltage (See App.(B.2)) MWIR and a cooled Mercury Cadmium Telluride (HgCdTe) LWIR camera are used for the dual band case. The spectral range of the MWIR and LWIR cameras are from $1.5\mu m$ to $5.1\mu m$ and from $7.7\mu m$ to $11.9\mu m$, respectively. A high pass filter with cut-on frequency at $3\mu m$ and a $8-12\mu m$ band pass filter are used in the respective MWIR and LWIR cameras. The horizontal and vertical frame sizes are 320 by 256 for both cameras. Moreover, $200mm$ lenses are used such that FOV_h and FOV_v values are 2.75 by 2.2 degrees. Both cameras have 14 bit resolution in the DN values. Spectral transfer functions of the MWIR and LWIR cameras are shown in Fig.7.1(a) and Fig.7.1(b), respectively.

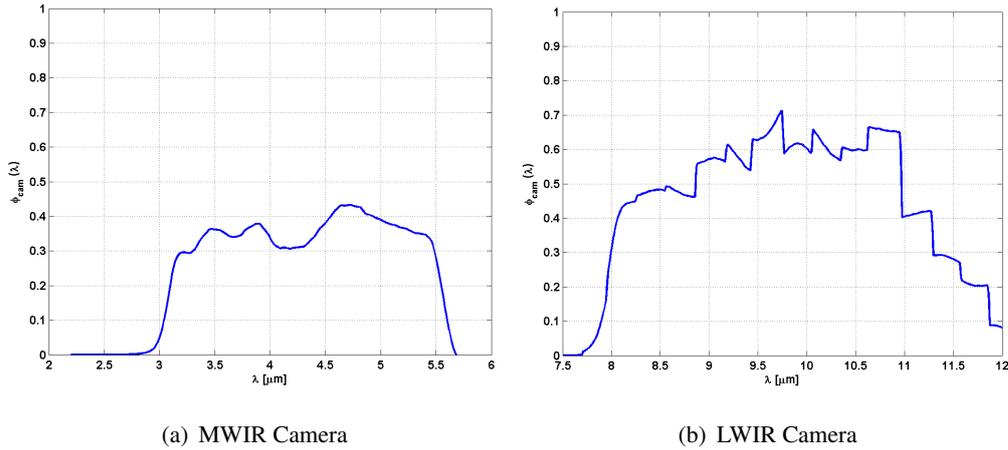


Figure 7.1: Spectral Transfer Functions of the Dual Band Measurement Setup

The cameras are located and aligned so that the need for registration is kept at minimum. In addition, to guarantee that the captured frames are synchronous in time, both cameras are triggered with the same square pulse signal generated by a signal generator.

The GPS coordinates of the cameras and also a predetermined location in the scene are measured using a hand-held GPS receiver.

In addition, to accurately estimate the atmospheric profile; air temperature, relative humidity, and pressure are measured using a meteorological data measurement system during the recordings for the dual band measurement system. For the single band measurement system; air temperature, relative humidity, and pressure are obtained from the measurements of [132]. For the visible scene, the measurement data are used directly in RT model. To construct the upper and lower atmospheric layers in RT model except the visible scene, the meteorological data are converted to sea level parameters using the models in [133, 134] and scaled accordingly. As expected, the GPS coordinates of the system are also taken into account in the construction of the layers.

CHAPTER 8

ANALYSIS RESULTS

Analysis results corresponding to each component of the proposed system are given in the following sections. Two sets of measurements are taken for each measurement setup. The first set of measurements are the ones where the scene is relatively close to the camera(s) and called near scenario. The second set of the measurements are the ones where the scene is relatively far away from the camera(s) and called far scenario.

8.1 Registration Results

Clearly, registration analyses have been made only for the dual band measurement setup. For a sample frame of the near scenario, the base image image is shown in Fig.8.1(a). In this figure and also for the rest of the thesis, in the gray scaled image visualizations the maximum and minimum values are mapped to the colors white and black, respectively. Hence, in-between values are gray scaled due to the linear mapping. The unregistered image and the registered image are shown in Fig.8.1(b) and Fig.8.1(c), respectively. Pixel indices of the base and the unregistered images differ for the same object. However, if the base and registered images are examined closely, it is seen that the pixel indices of both images correspond to almost the same objects in the scene.

Similarly for a sample frame of the far scenario, the base image image is shown in Fig.8.2(a). The unregistered image and the registered image are shown in Fig.8.2(b) and Fig.8.2(c), respectively. Although registration problem is harder in the far scenario case than the near scenario case, the affine registration algorithm yields satis-

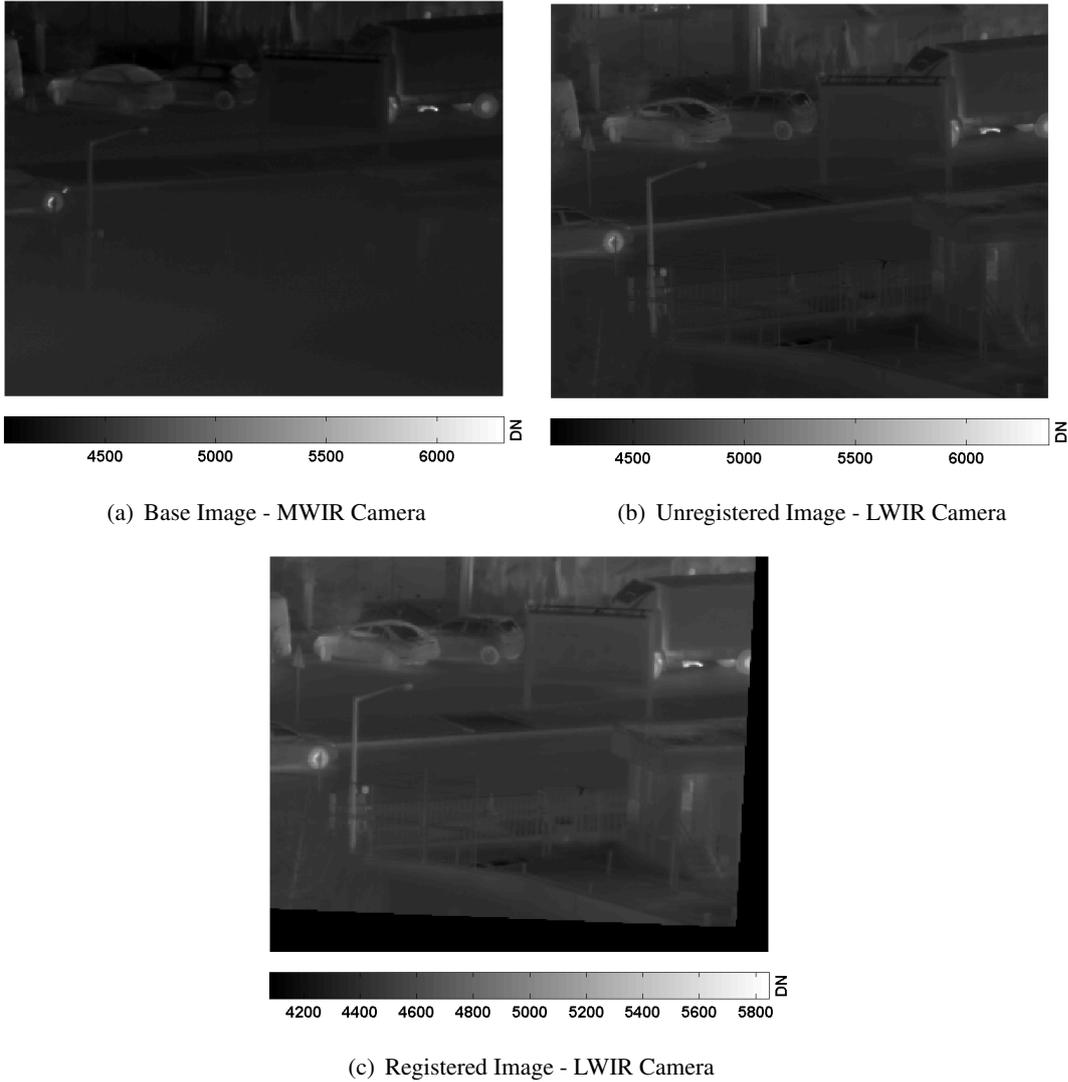


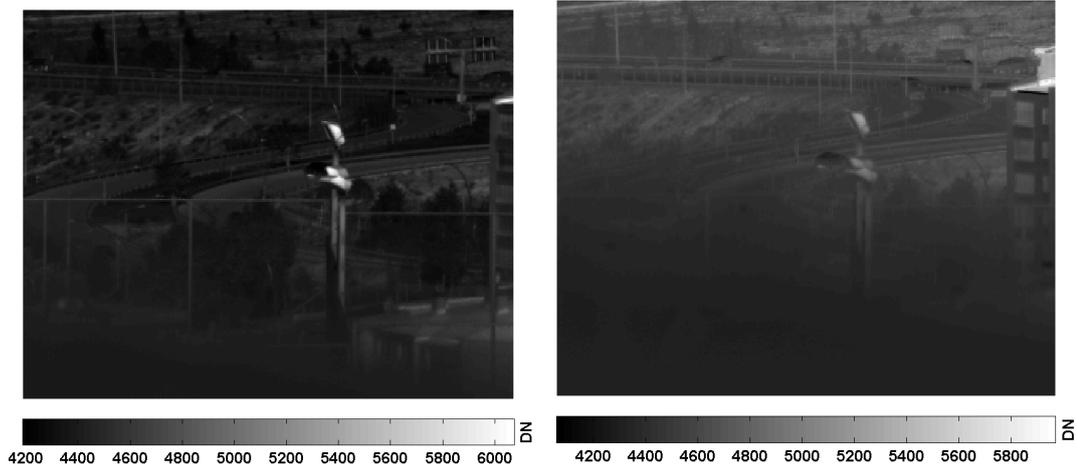
Figure 8.1: Dual Band Measurement Setup Registration Results - Near Scenario

factory results especially in the middle part of the image.

8.2 Geometry Estimation Results

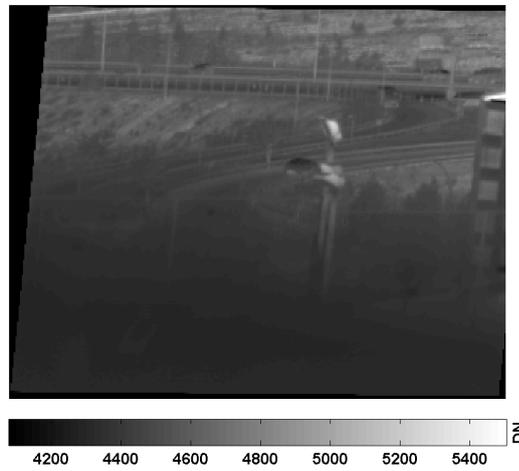
8.2.1 Geometry Estimation Results for the Single Band Measurements

Two sets of measurements are taken with the single band measurement setup. The output of the geometry estimation process for the near scenario, which are the range and zenith angle values for each pixel, are shown in Fig.8.3(a) and Fig.8.3(b), respectively. The camera altitude in this case is $h = 985.97m$. The range values in the



(a) Base Image - MWIR Camera

(b) Unregistered Image - LWIR Camera

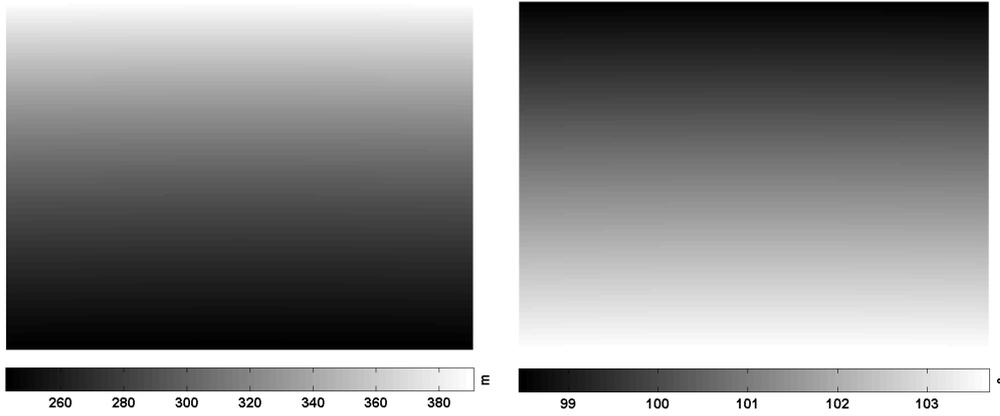


(c) Registered Image - LWIR Camera

Figure 8.2: Dual Band Measurement Setup Registration Results - Far Scenario

frame are estimated to be in the interval from $250m$ to $390m$. Corresponding zenith angles are in 104° to 98° interval.

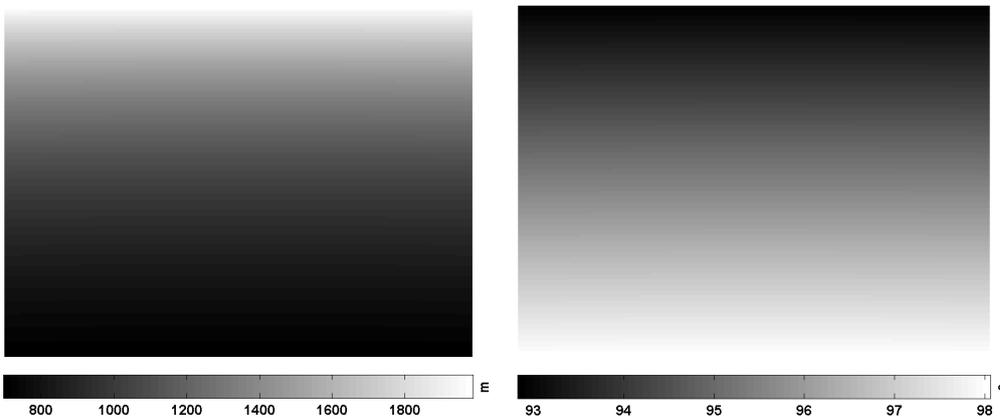
The output of the geometry estimation process for the far scenario, which are the range and zenith angle values for each pixel, are shown in Fig.8.4(a) and Fig.8.4(b), respectively. The camera altitude is again $h = 985.97m$. As expected the range values are larger in the far scenario, which are in $750m$ to $2km$ interval. In addition, corresponding zenith angles are from 98° to 93° .



(a) Range Values

(b) Zenith Angle Values

Figure 8.3: Single Band Measurement Setup Geometry Estimation Results - Near Scenario



(a) Range Values

(b) Zenith Angle Values

Figure 8.4: Single Band Measurement Setup Geometry Estimation Results - Far Scenario

8.2.2 Geometry Estimation Results for the Dual Band Measurements

Two sets of measurements are taken with the dual band measurement setup. The output of the geometry estimation process for the near scenario, which are the range and zenith angle values of each pixel, are shown in Fig.8.5(a) and Fig.8.5(b), respectively for the MWIR camera. The range and zenith angle values for the LWIR camera are shown in Fig.8.5(c) and Fig.8.5(d), respectively. The camera altitude is $h = 985.97m$ for both of the cameras. Range and zenith angle values across MWIR and LWIR cameras are very close to each other since the cameras are located side by side and oriented towards almost the same area in the scene. Range values are from $270m$

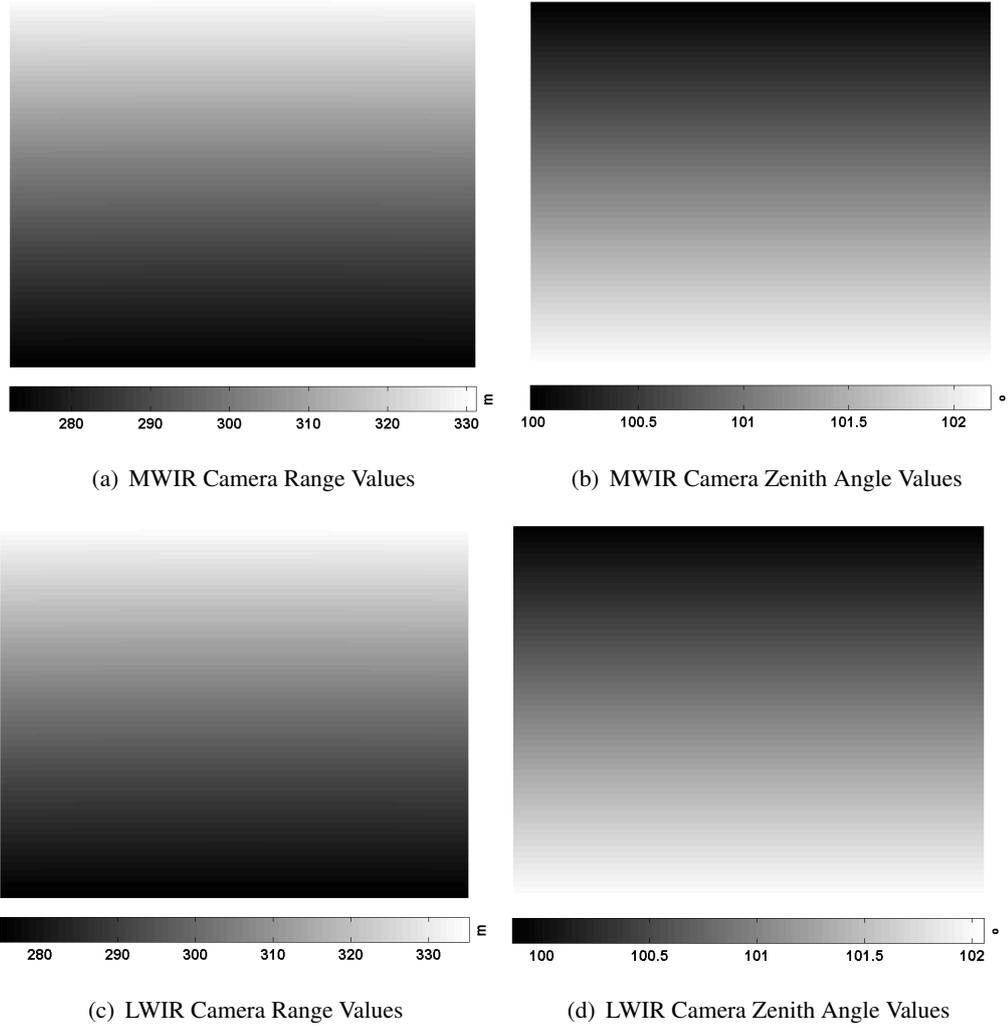


Figure 8.5: Dual Band Measurement Setup Geometry Estimation Results - Near Scenario

to $335m$. Corresponding zenith angles are from 102° to 100° . Clearly, these intervals are narrower compared to the single band case since the FOV values of the dual band measurement system is smaller than the FOV value of the single band system. Moreover, since the resolution of the standard GPS receivers is on the order of a few meters, the cameras are assumed to have the same position.

The output of the geometry estimation process for the far scenario, which are the range and zenith angle values for each pixel, are shown in Fig.8.6(a) and Fig.8.6(b), respectively for the MWIR camera. The range and zenith angle values for the LWIR camera are shown in Fig.8.6(c) and Fig.8.6(d), respectively. The camera altitude is $h = 985.97m$ for both of the cameras. Range values are estimated to be in the interval

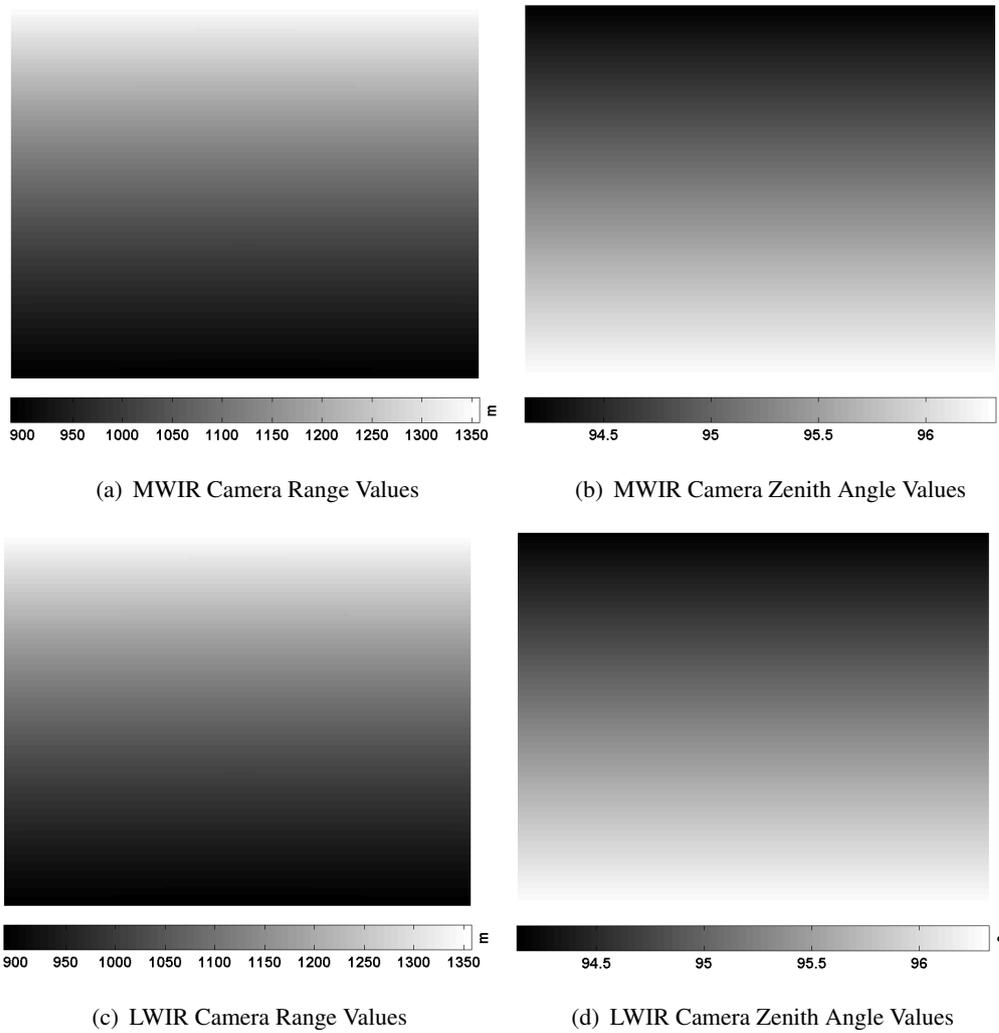


Figure 8.6: Dual Band Measurement Setup Geometry Estimation Results - Far Scenario

from 900m to 1350m. Corresponding zenith angles are in the interval from 96.5° to 94°. Again, range and zenith angle values are very close to each other for MWIR and LWIR cameras. Moreover, the intervals are narrower compared to the far scenario measurements of the single band system.

8.3 Atmospheric Effects Removal Results

In the atmospheric effects removal, Riemann sum is used to numerically compute the respective continuous integral expressions. In the related equations the wavelength step size, $\Delta\lambda$, is used to compute the Riemann sums. $\Delta\lambda$ value is obtained according

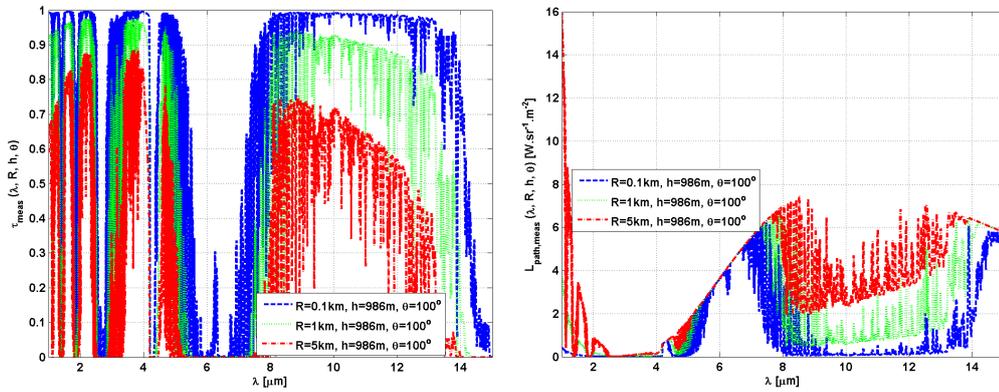
to,

$$\Delta\lambda = \frac{(\lambda_{high} - \lambda_{low})}{999} \quad (8.1)$$

which means that the spectral computations are done at 1000 points.

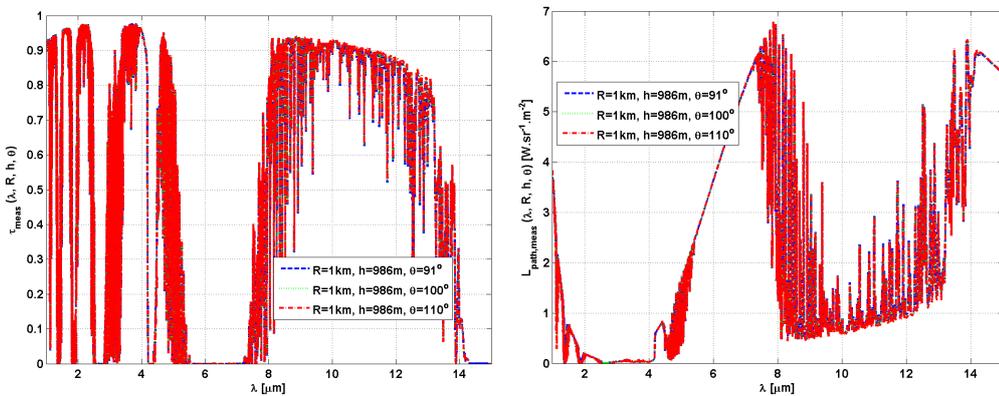
8.3.1 Radiative Transfer Computations Results

RT computations for the output values of the geometry estimation are made using Modtran[®]. The effect of the range on the spectral transmittance and the path radiance are shown in Fig.8.7(a) and Fig.8.7(b), respectively by using sample Modtran[®] outputs. Moreover, the effect of the zenith angle on the spectral transmittance and the



(a) Spectral Transmittance (b) Spectral Path Radiance
Figure 8.7: Effect of the Range on the Radiative Transfer

path radiance are shown in Fig.8.8(a) and Fig.8.8(b), respectively. The range value



(a) Spectral Transmittance (b) Spectral Path Radiance
Figure 8.8: Effect of the Zenith Angle on the Radiative Transfer

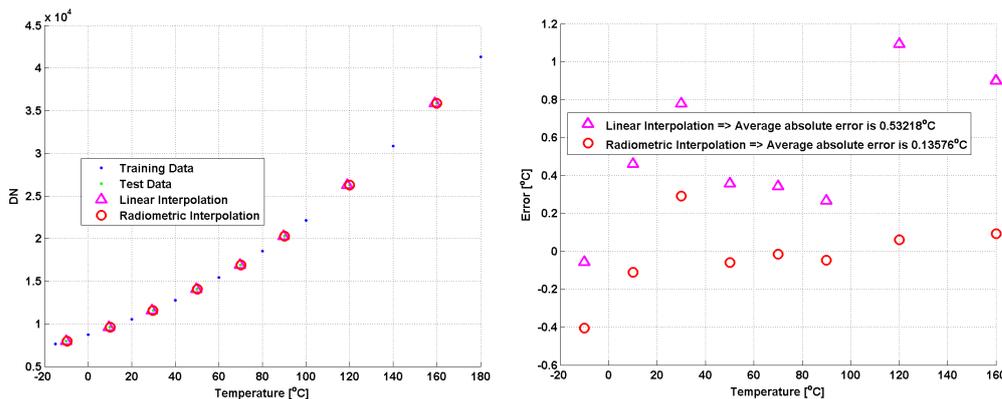
variations have a serious effect on the transmittance and path radiance for the measurement scenarios. However, zenith angle variations seem to have a very small effect on the RT model output. This is because our measurements are taken with lenses that have small FOV values. For geometries such as $\theta \in [100^\circ, 140^\circ]$, zenith angle has considerable impacts on both transmittance and path radiance, which is the case for cameras with larger FOV values.

8.3.2 Temperature Calibration with Radiometric Interpolation Results

8.3.2.1 Temperature Calibration with Radiometric Interpolation Results for the Single Band Measurements

For the analyses, temperature calibration of the camera is a critical step. Therefore, firstly the validity of the proposed calibration method is investigated. In this context, a calibration with 17 points is done for the LWIR camera of the single band measurement setup. Half of these calibration points are selected as training pairs. The rest of the points are used to test the reference linear interpolation and the proposed radiometric interpolation methods.

The training and test pairs together with the results of the linear and radiometric interpolation methods are shown in Fig.8.9(a). The calibration error that is the tem-



(a) Temperature Calibration

(b) Calibration Error

Figure 8.9: Single Band Measurement Setup - Justification of the Temperature Calibration Method

perature difference between the test point and the result of each method are shown

in Fig.8.9(b). The average absolute error of the methods are shown on the legend of the figure, as well. The average absolute error of the linear interpolation is $532mK$, whereas it is $136mK$ for the proposed method. The typical NETD value of the camera that is used in these measurements, is specified to be $80mK$ at $30^{\circ}C$. Therefore, the results show that the proposed method performs better and yields calibration error values close to the NETD of the camera. Moreover, once we use all the calibration pairs rather than half of the data, the error performance is expected to improve further.

Temperature calibration with linear interpolation and also the proposed calibration method using all the calibration pairs are shown in Fig.8.10(a). The temperature dif-

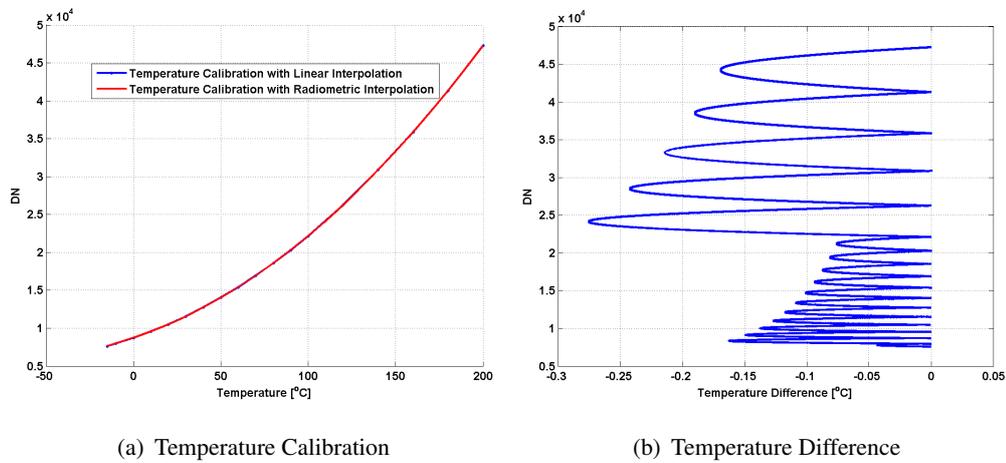


Figure 8.10: Single Band Measurement Setup - Radiometric Calibration

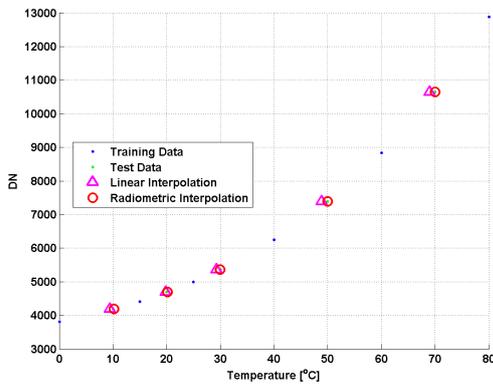
ference between the reference and the proposed methods is also given in Fig.8.10(b). Clearly, the temperature difference values are equal to zero at the calibration points. Besides, it is observed that the linear interpolation method always overestimates the temperature values in between the calibration points. This is due to the slope of the Planck curve for the temperature and pass-band of the camera.

8.3.2.2 Temperature Calibration with Radiometric Interpolation Results for the Dual Band Measurements

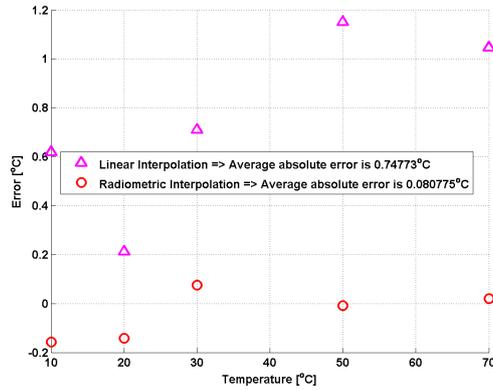
The same analysis is repeated for the dual band measurement setup. The validity of the proposed calibration method for both cameras is investigated. In this context, a calibration with 11 and a calibration with 17 points are done for the MWIR and LWIR cameras, respectively. Half of these calibration points are selected as training

pairs. The rest of the points are used to test the reference linear interpolation and the proposed radiometric interpolation.

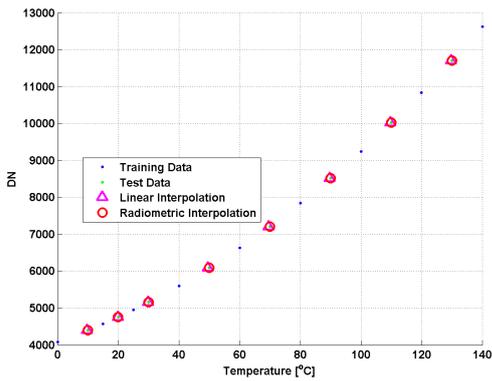
The training and test pairs together with the results of the linear and radiometric interpolation methods are shown in Fig.8.11(a) for the MWIR camera. The calibration



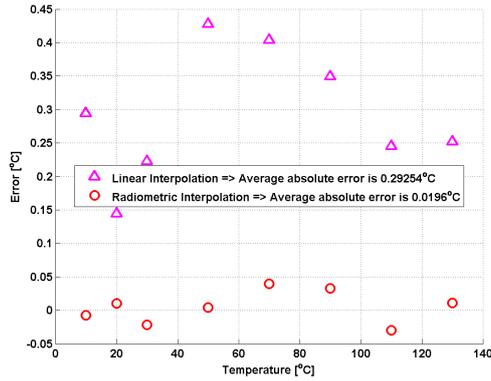
(a) Temperature Calibration - MWIR Camera



(b) Calibration Error - MWIR Camera



(c) Temperature Calibration - LWIR Camera



(d) Calibration Error - LWIR Camera

Figure 8.11: Dual Band Measurement Setup - Justification of the Temperature Calibration Method

error that is the temperature difference between the test point and the result of each method are shown in Fig.8.11(b) for the MWIR camera. Similarly, the calibration and the error graphs for the LWIR camera are shown in Fig.8.11(c) and in Fig.8.11(d), respectively. The average absolute error of the methods are shown on the legend of the figures, as well. For the MWIR camera, the average absolute error of the linear interpolation is $748mK$, whereas it is $80.8mK$ for the proposed method. For the LWIR camera, the respective values are $292mK$ and $19.6mK$. The typical NETD value of both MWIR and LWIR cameras are specified to be $20mK$ at $25^{\circ}C$.

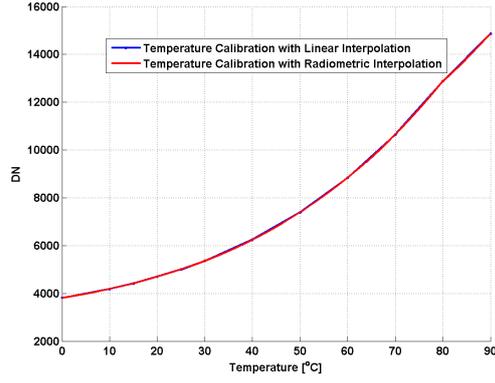
For the LWIR camera, the proposed method performs almost perfectly. This is due to the fact that the spectral variations in the LWIR band is smoother than the variations in the MWIR band. On the other hand, for the MWIR camera the proposed method also performs much better than the linear interpolation method but the average error value of the proposed method is slightly higher than the NETD value. However, once we use all the calibration pairs rather than half of the data, the error performance tends to improve. Moreover, $80.8mK$ is so small that it is negligible in the analyses. Hence, it is shown that the proposed radiometric interpolation method yields very efficient results. In addition, both the performance and the performance gain is better for the dual band cameras than the single band camera. This is because, single band system camera has a non-cooled micro bolometer detector. However, the cameras of the dual band system have cooled detectors, which means they are much more sensitive measurement devices.

Temperature calibration with linear interpolation and also the proposed calibration method using all the calibration pairs are shown in Fig.8.12(a) for the MWIR camera. The temperature difference between the reference and the proposed methods is also given in Fig.8.12(b) for the MWIR camera. Similarly, calibration and the difference graphs for the LWIR camera are shown in Fig.8.12(c) and in Fig.8.12(d), respectively.

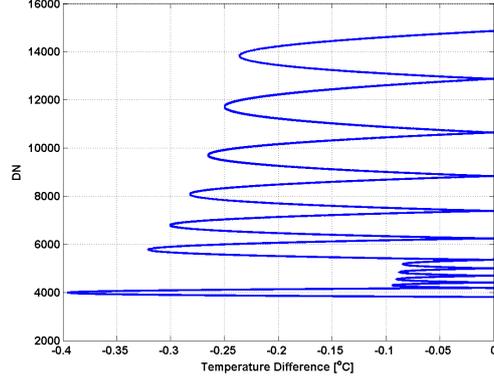
Although the dynamic range of the MWIR camera is less than the dynamic range of the LWIR camera, the difference between the two methods is larger in the MWIR band. This is mainly because the spectral variations in the MWIR band is greater than the variations in the LWIR band, especially due to the molecular absorptions.

8.3.3 Atmospheric Effects Removal Results for the Single Band Measurements

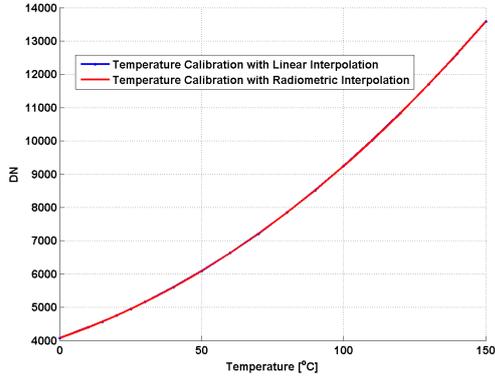
A sample frame of the near scenario is analyzed. The DN image captured by the LWIR camera and the result of the temperature calibration for the respective images are shown in Fig.8.13(a) and Fig.8.13(b), respectively. According to the radiometric calibration, the contents of the scene in terms of calibration temperature are in the interval approximately from $25^{\circ}C$ to $62^{\circ}C$. The second term of the left hand side of Eq.(3.32), which is the weighted path radiance component is shown in Fig.8.13(c). In addition, τ_w is shown in Fig.8.13(d).



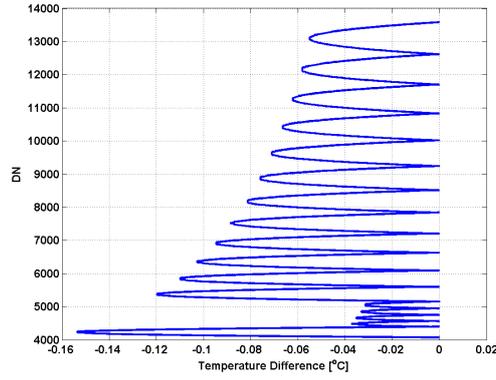
(a) Temperature Calibration - MWIR Camera



(b) Temperature Difference - MWIR Camera



(c) Temperature Calibration - LWIR Camera



(d) Temperature Difference - LWIR Camera

Figure 8.12: Dual Band Measurement Setup - Radiometric Calibration

The apparent radiance that is the left hand side of Eq.(3.32) is shown in Fig.8.14(a). The distribution of the apparent radiance image is shown in Fig.8.14(b). The estimated source radiance computed using Eq.(3.34) and the image distribution are given in Fig.8.14(c) and Fig.8.14(d), respectively. As clearly seen, the distribution of the image is changed due to the removed atmospheric effects. Moreover, the blackbody radiance image, which is the radiance computed with the assumption of a blackbody object ($\epsilon_{x=meas}(i, j, \lambda) = 1 \quad \forall i, j, \lambda$) in an ideal atmosphere ($\tau_{x=meas}(i, j, \lambda) = 1 \quad \forall i, j, \lambda$) with an ideal camera ($\varphi_{cam}(\lambda) = 1 \quad \forall \lambda$), is given in Fig.8.14(e). The distribution of the blackbody image is shown in Fig.8.14(f).

A sample frame of the far scenario is also analyzed. The DN image captured by the LWIR camera and the result of the temperature calibration for the respective images are shown in Fig.8.15(a) and Fig.8.15(b), respectively. The contents of the scene in terms of calibration temperature are in the interval approximately from $20^\circ C$ to $60^\circ C$. The weighted path radiance component and also τ_w are shown in Fig.8.15(c) and in

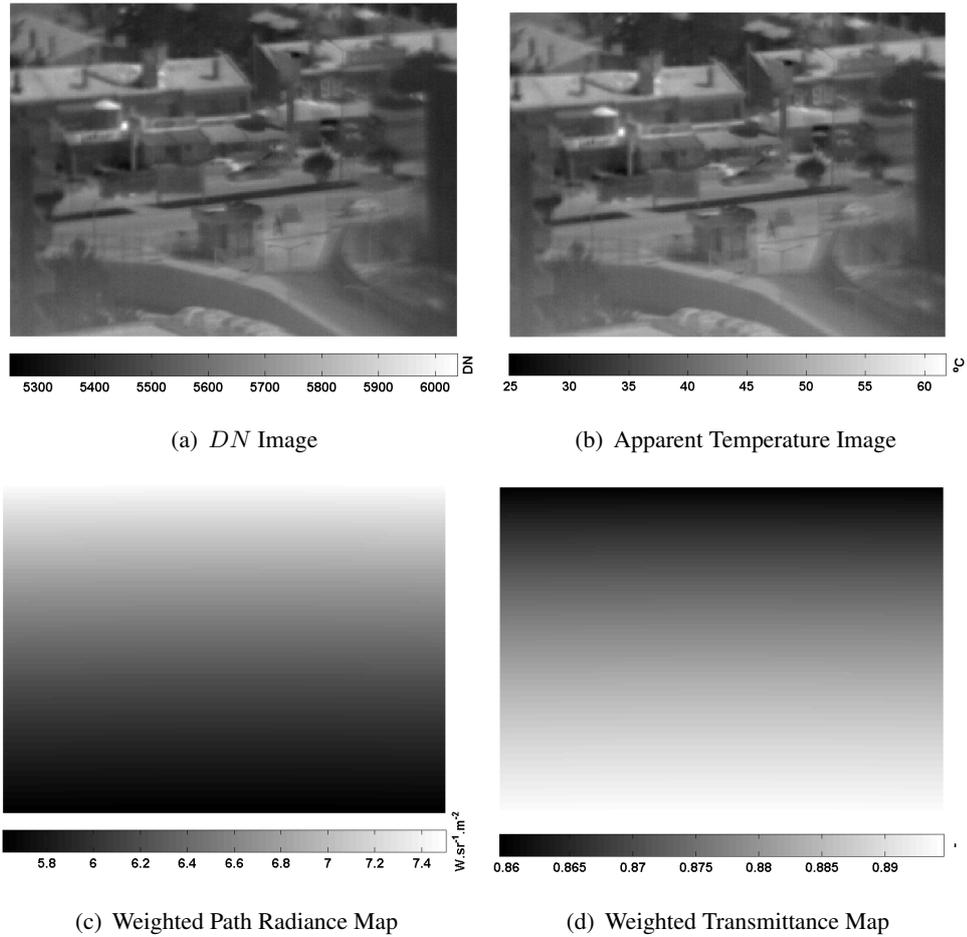
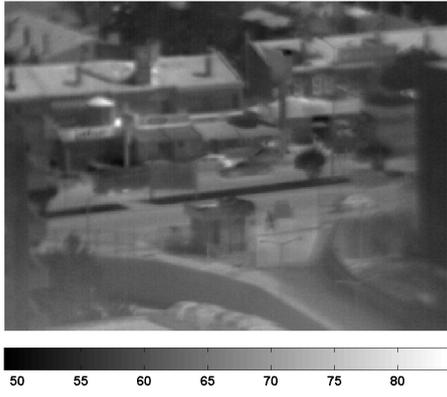


Figure 8.13: Single Band Measurement Setup Atmospheric Effects Removal Analysis Results - Near Scenario

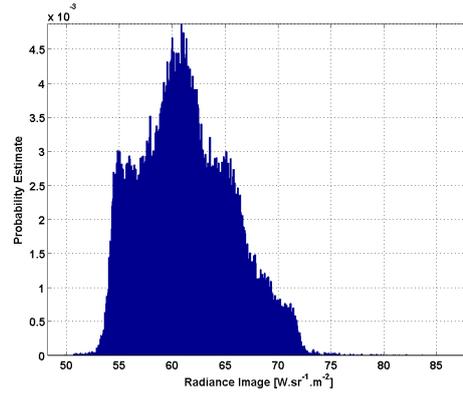
Fig.8.15(d), respectively.

The apparent radiance image is shown in Fig.8.16(a). The distribution of the apparent radiance image is shown in Fig.8.16(b). The estimated source radiance and the image distribution are given in Fig.8.16(c) and Fig.8.16(d), respectively. Similar to the near scenario, the distribution of the image is changed due to the atmospheric correction. Moreover, the blackbody radiance image is given in Fig.8.16(e). The distribution of the blackbody image is shown in Fig.8.16(f).

In addition to the single frame analyses, a global analysis has also been made for all the recordings of the single band measurements. Among these measurements first measurements in time are the far scenarios, whereas the second measurements are the near scenarios. For the far scenarios, there are 11 recordings. For each recording a ROI, which is 9×8 pixels wide, is specified. The ROI is a region corresponding



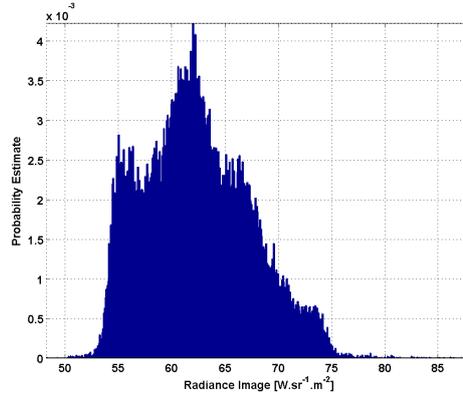
(a) Apparent Radiance Image



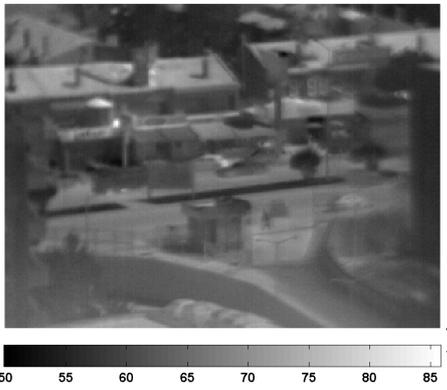
(b) Apparent Radiance Image Distribution



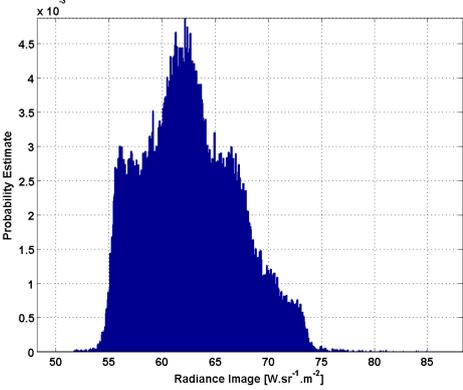
(c) Estimated Source Radiance Image



(d) Estimated Source Radiance Image Distribution



(e) Blackbody Radiance Image



(f) Blackbody Radiance Image Distribution

Figure 8.14: Single Band Measurement Setup Atmospheric Effects Removal Radiance Images - Near Scenario

only to the asphalt in the scene. In these recordings 100 frames are determined such that there is no other object occluding the asphalt. For the near scenarios, there are 6 recordings. For each recording a ROI, which is 31×4 pixels wide, is specified.

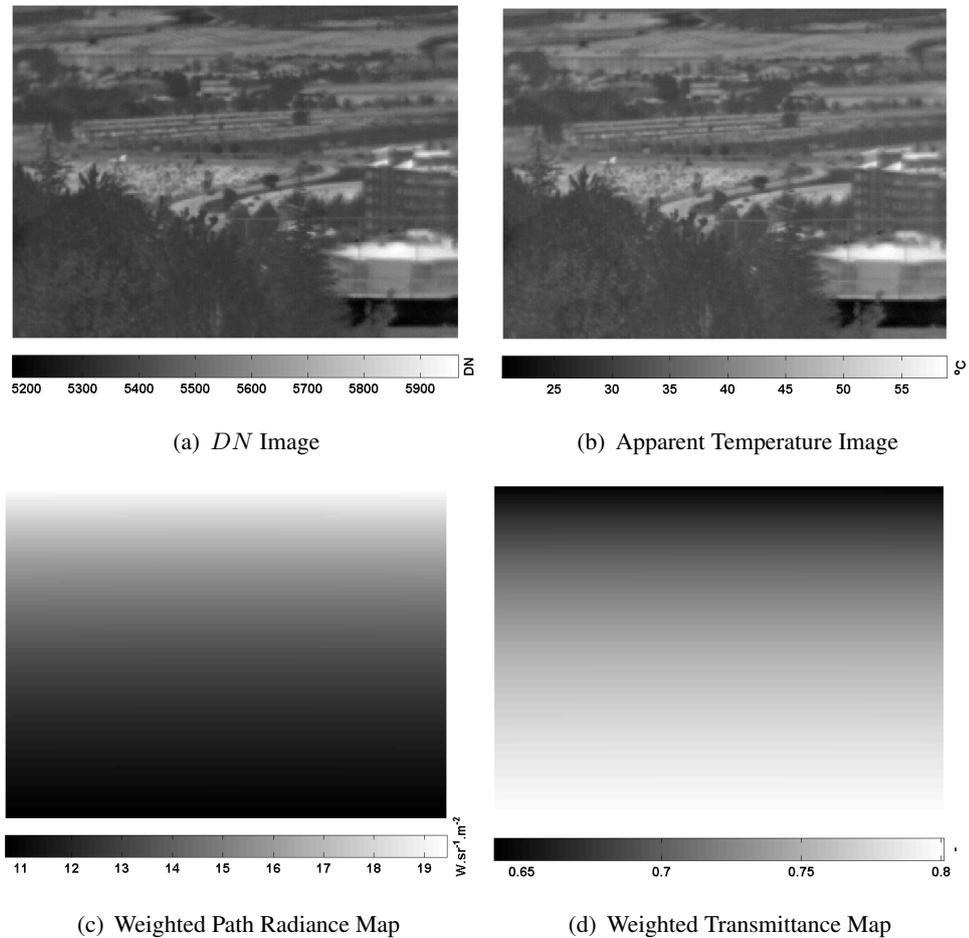
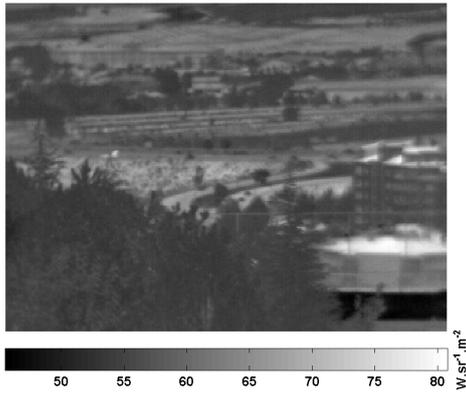


Figure 8.15: Single Band Measurement Setup Atmospheric Effects Removal Analysis Results - Far Scenario

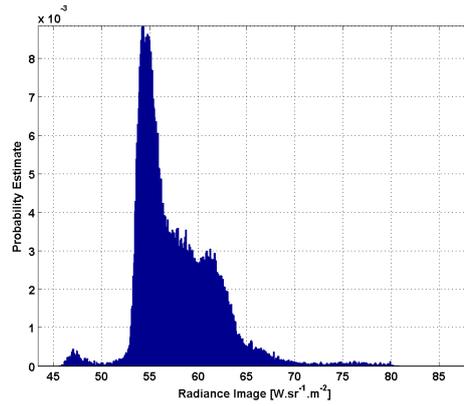
The ROI is again a region corresponding only to the asphalt in the scene. In these recordings 100 frames are determined such that there is no other object occluding the asphalt. Then the apparent radiance, blackbody radiance, and also the estimated source radiance values for each of these pixels in the determined frames are computed similar to the single frame analysis. The temporal and spatial average and also the standard deviation of each scenario is computed and shown as an error bar plot in Fig.8.17.

The apparent radiance values are similar to the blackbody and estimated source radiance values. This is because, the system transfer function of the camera is not available, so it is assumed to be ideal within the pass-band of the camera.

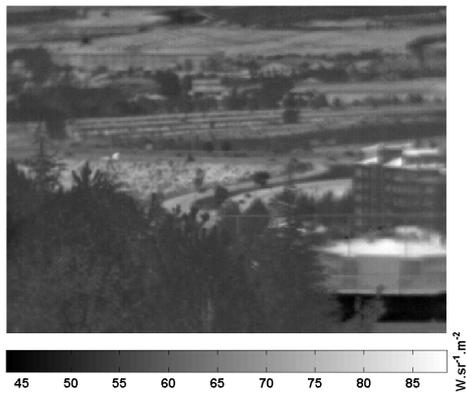
ROIs are chosen such that they contain only asphalt because according to spectral library in [135], asphalt behaves almost like a blackbody. The spectral directional



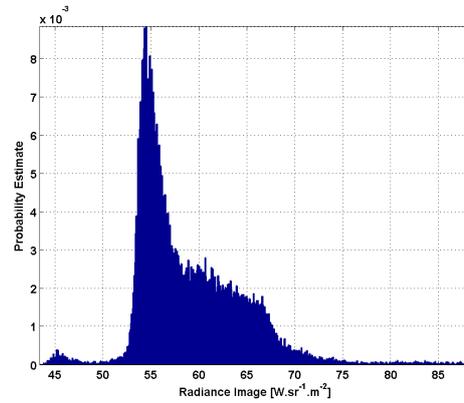
(a) Apparent Radiance Image



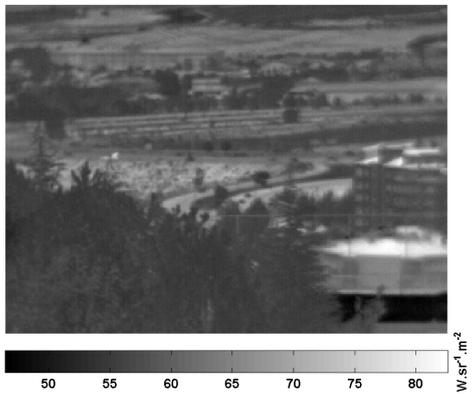
(b) Apparent Radiance Image Distribution



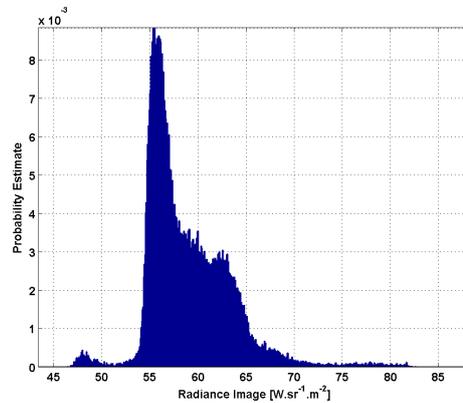
(c) Estimated Source Radiance Image



(d) Estimated Source Radiance Image Distribution



(e) Blackbody Radiance Image



(f) Blackbody Radiance Image Distribution

Figure 8.16: Single Band Measurement Setup Atmospheric Effects Removal Radiance Images - Far Scenario

hemispherical reflectance of the asphalt in the spectral pass-band of the camera is given in Fig.8.18. The average directional hemispherical reflectance in the band of interest is approximately 0.05 and the spectral deviation is around 0.015. Therefore,

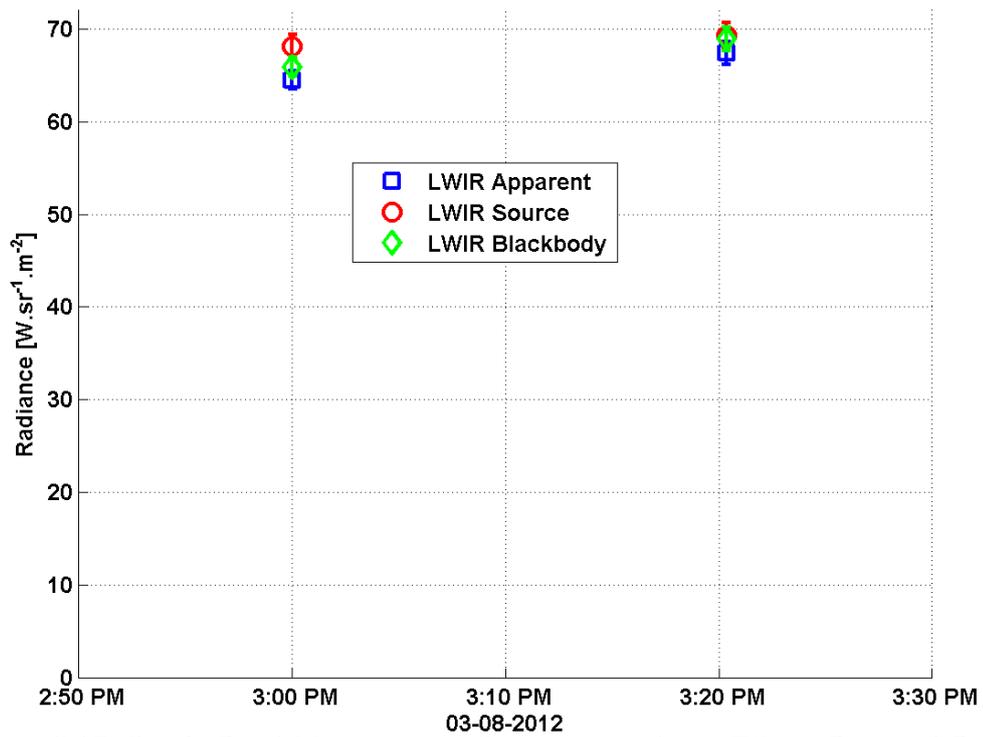


Figure 8.17: Single Band Measurement Setup Atmospheric Effects Removal Results

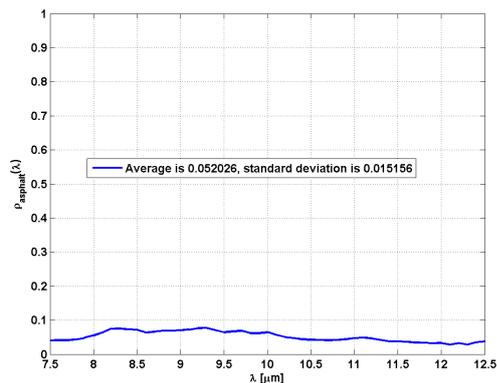


Figure 8.18: Spectral Reflectance of the Asphalt for the Single Band Measurement Setup

reflections from asphalt surface can be neglected. Moreover, the air temperature measurements are the same for the respective times. Hence, the dominant difference between the blackbody radiance and the estimated source radiance values are considered to be because of the atmospheric effects. In other words, if there were an ideal atmosphere; the blackbody and estimated source radiance values are expected to have almost the same values.

The difference between the two times is minimum for the estimated source radiance since the atmospheric effects are removed. The remaining difference is due to the

solar energy absorption of the asphalt. Therefore, with the help of atmospheric correction method, the variation of the radiance between the recordings corresponding to different times is decreased as expected.

8.3.4 Atmospheric Effects Removal Results for the Dual Band Measurements

A sample frame of the near scenario for the MWIR camera is analyzed. The DN image captured by the MWIR camera and the result of the temperature calibration for the respective image are shown in Fig.8.19(a) and Fig.8.19(b), respectively. According to

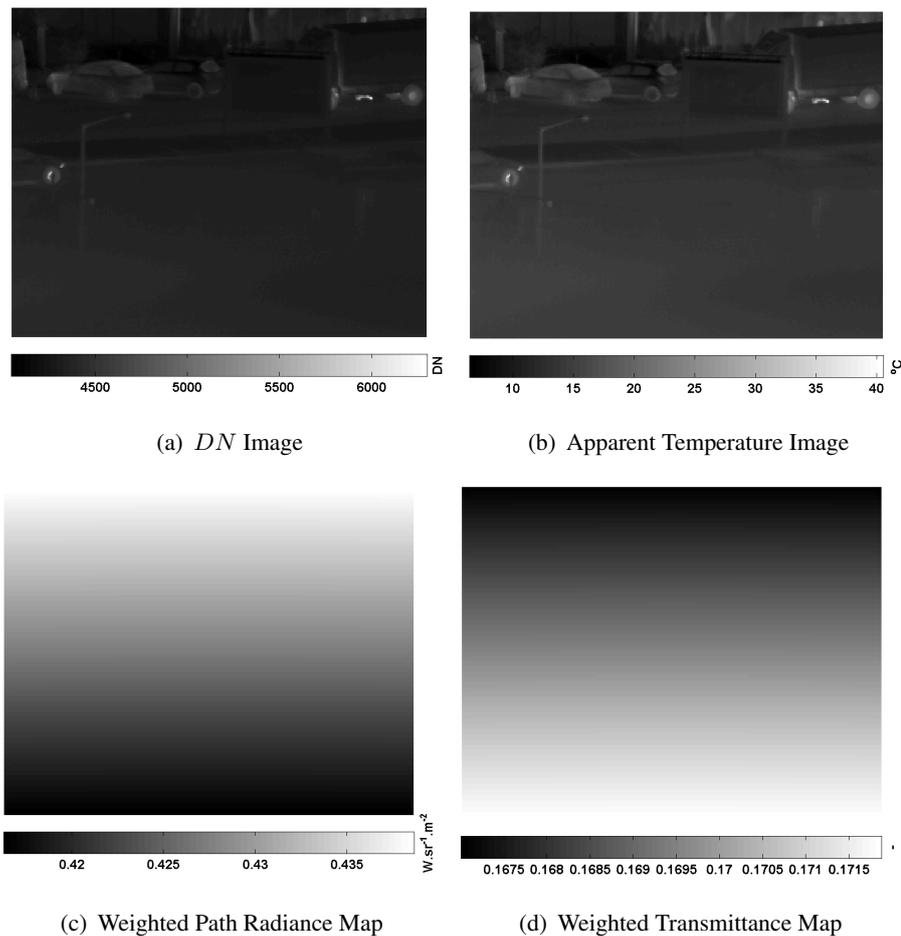
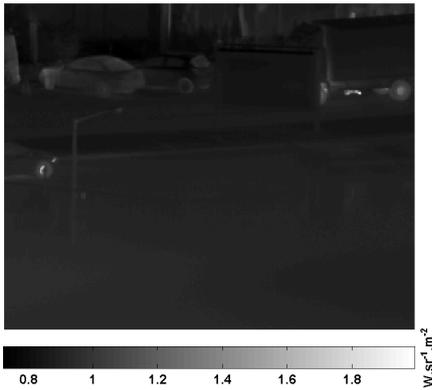


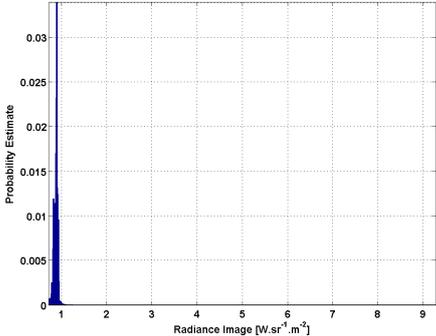
Figure 8.19: Dual Band Measurement Setup Atmospheric Effects Removal Analysis Results - MWIR Camera - Near Scenario

the radiometric calibration, the contents of the scene in terms of calibration temperature are in the interval approximately from $5^{\circ}C$ to $41^{\circ}C$. The weighted path radiance component for the MWIR camera is shown in Fig.8.19(c). In addition, τ_w for the MWIR camera is shown in Fig.8.19(d).

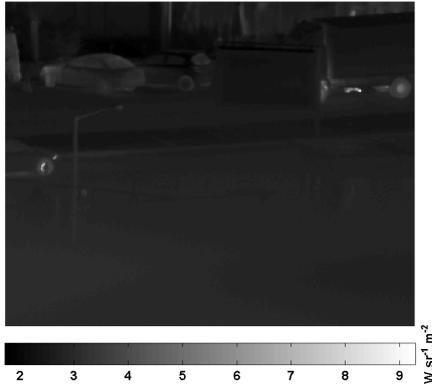
The apparent radiance is shown in Fig.8.20(a). The distribution of the apparent ra-



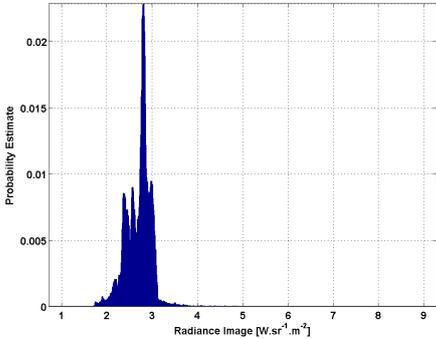
(a) Apparent Radiance Image



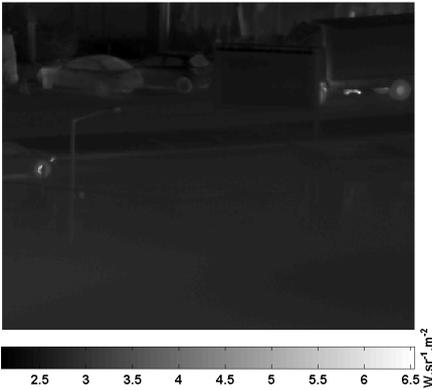
(b) Apparent Radiance Image Distribution



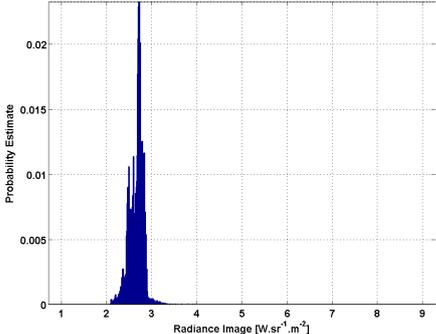
(c) Estimated Source Radiance Image



(d) Estimated Source Radiance Image Distribution



(e) Blackbody Radiance Image



(f) Blackbody Radiance Image Distribution

Figure 8.20: Dual Band Measurement Setup Atmospheric Effects Removal Radiance Images - MWIR Camera - Near Scenario

diance image is shown in Fig.8.20(b). The estimated source radiance and the image distribution are given in Fig.8.20(c) and Fig.8.20(d), respectively. As clearly seen, the

distribution of the image is changed due to the removed atmospheric effects. Moreover, the blackbody radiance image is given in Fig.8.20(e). The distribution of the blackbody image is shown in Fig.8.20(f).

A sample frame of the near scenario for the LWIR camera is analyzed, too. The *DN* image captured by the LWIR camera and the result of the temperature calibration for the respective image are shown in Fig.8.21(a) and Fig.8.21(b), respectively. The con-

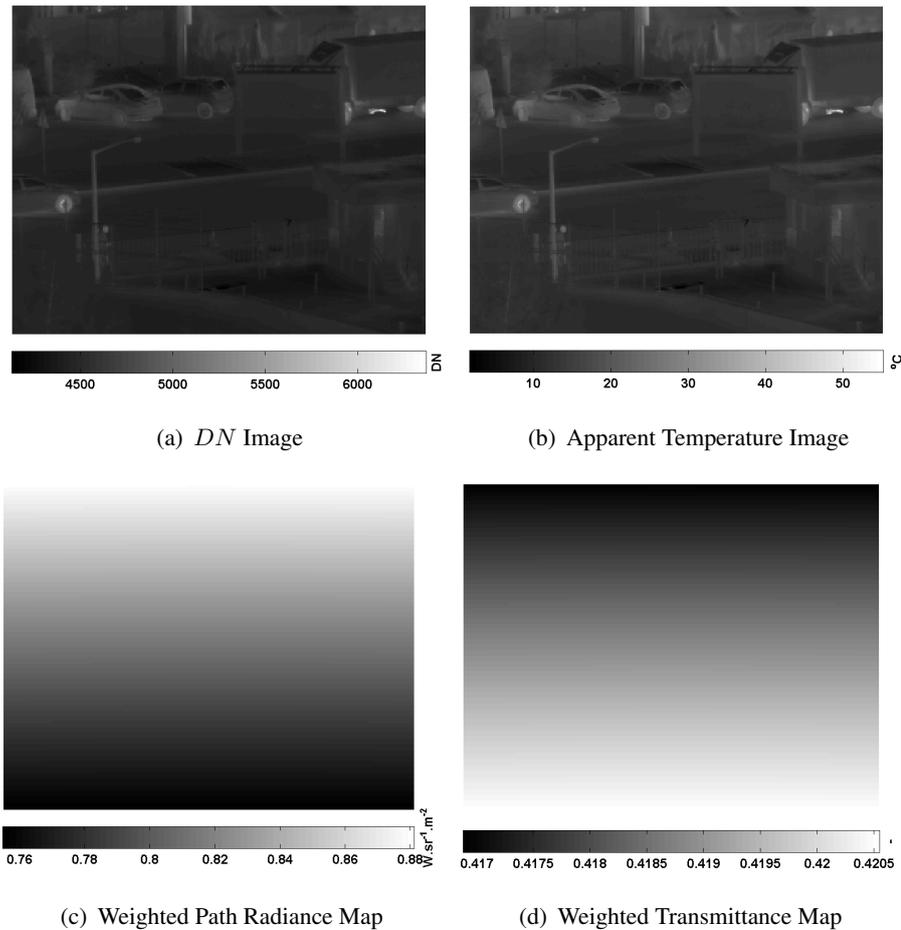
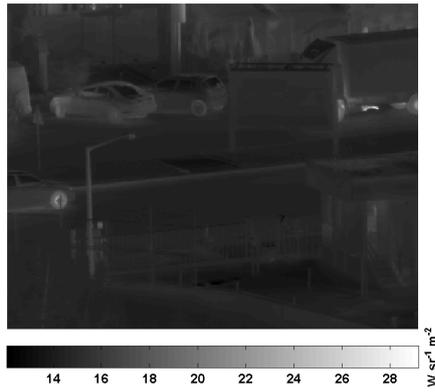


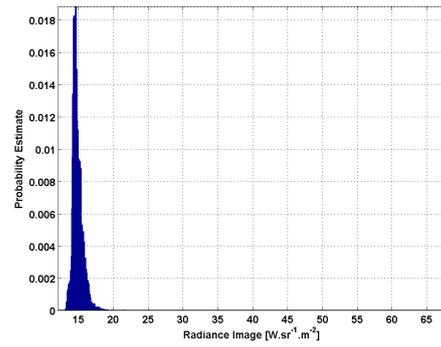
Figure 8.21: Dual Band Measurement Setup Atmospheric Effects Removal Analysis Results - LWIR Camera - Near Scenario

tents of the scene in terms of calibration temperature are in the interval approximately from $3^{\circ}C$ to $55^{\circ}C$. Since these values are not the surface temperatures but the apparent temperatures, the values of the MWIR and LWIR cameras are not similar. The weighted path radiance component for the LWIR camera is shown in Fig.8.21(c). In addition, τ_w for the LWIR camera is shown in Fig.8.21(d).

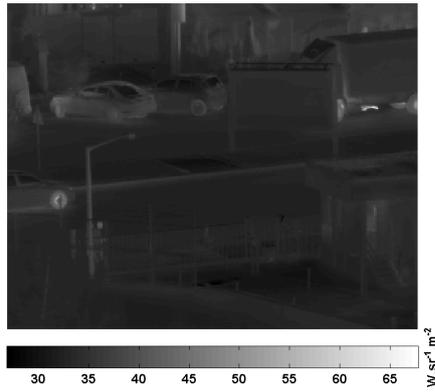
The apparent radiance image is shown in Fig.8.22(a). The distribution of the appar-



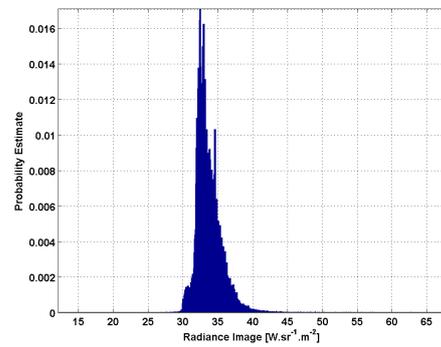
(a) Apparent Radiance Image



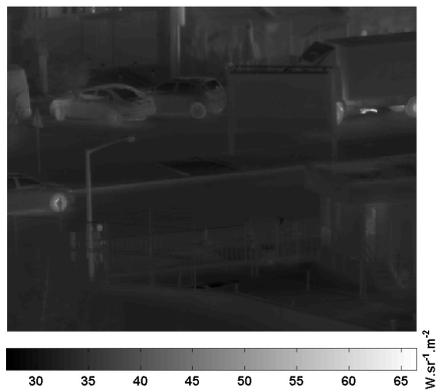
(b) Apparent Radiance Image Distribution



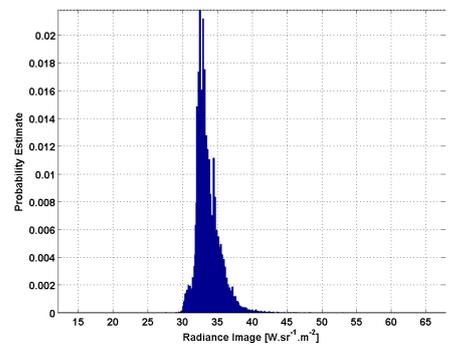
(c) Estimated Source Radiance Image



(d) Estimated Source Radiance Image Distribution



(e) Blackbody Radiance Image



(f) Blackbody Radiance Image Distribution

Figure 8.22: Dual Band Measurement Setup Atmospheric Effects Removal Radiance Images - LWIR Camera - Near Scenario

ent radiance image is shown in Fig.8.22(b). The estimated source radiance and the image distribution are given in Fig.8.22(c) and Fig.8.22(d), respectively. Similar to the MWIR camera case, the distribution of the image is changed due to the removed

atmospheric effects. Moreover, the blackbody radiance image is given in Fig.8.22(e). The distribution of the blackbody image is shown in Fig.8.22(f).

A sample frame of the far scenario for the MWIR camera is also analyzed. The DN image captured by the MWIR camera and the result of the temperature calibration for the respective image are shown in Fig.8.23(a) and Fig.8.23(b), respectively. The con-

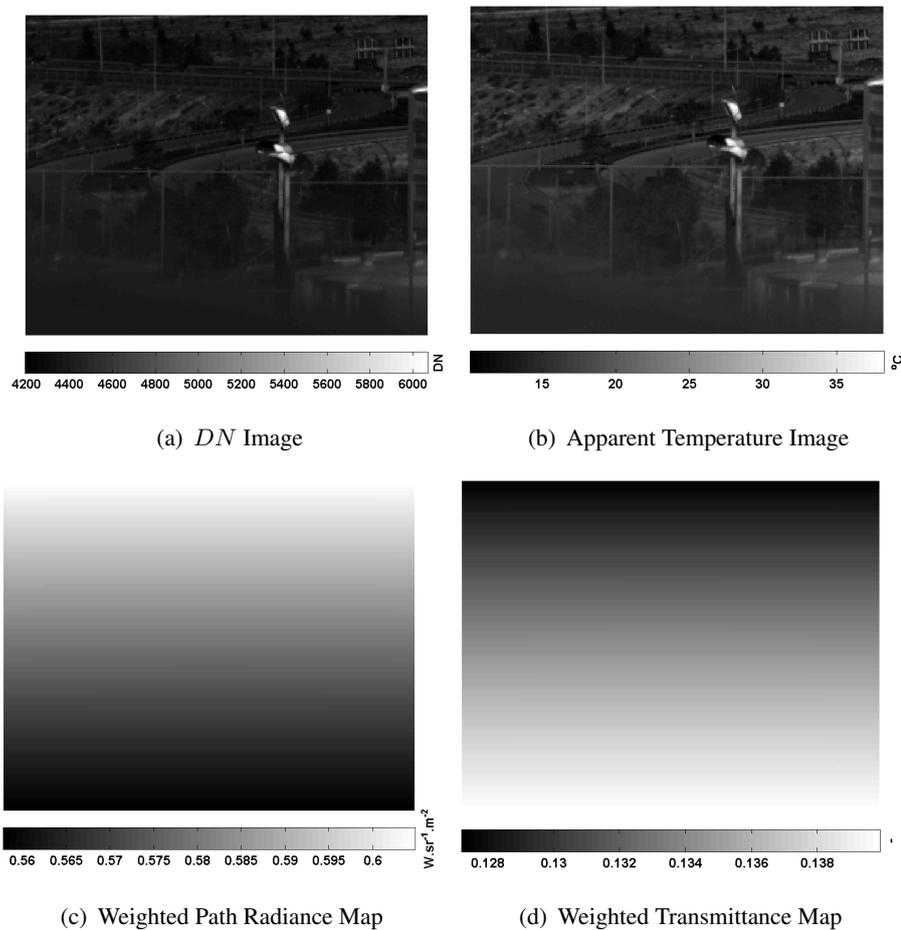
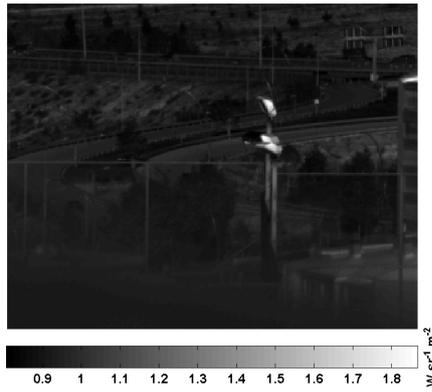


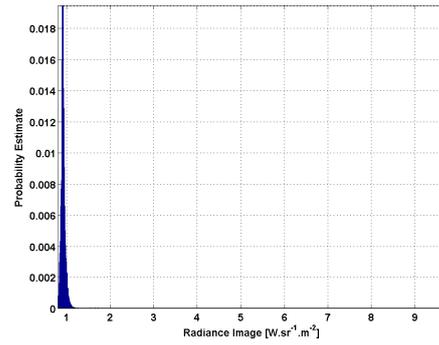
Figure 8.23: Dual Band Measurement Setup Atmospheric Effects Removal Analysis Results - MWIR Camera - Far Scenario

tents of the scene in terms of calibration temperature are in the interval approximately from $10^{\circ}C$ to $39^{\circ}C$. The weighted path radiance component for the MWIR camera is shown in Fig.8.23(c). In addition, τ_w for the MWIR camera is shown in Fig.8.23(d).

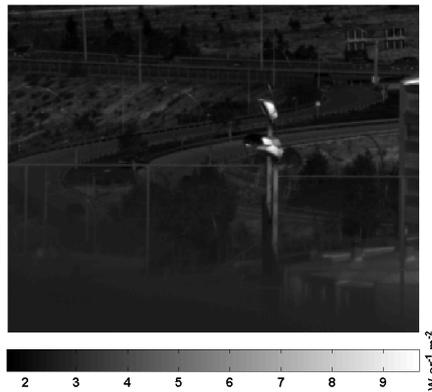
The apparent radiance is shown in Fig.8.24(a). The distribution of the apparent radiance image is shown in Fig.8.24(b). The estimated source radiance and the image distribution are given in Fig.8.24(c) and in Fig.8.24(d), respectively. As clearly seen, the distribution of the image is changed due to the removed atmospheric effects.



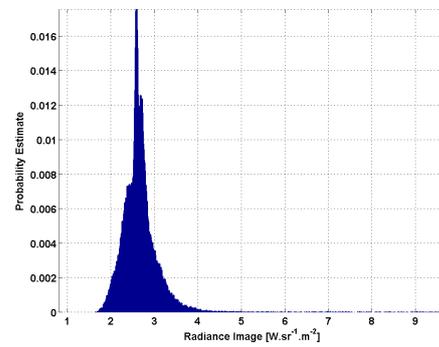
(a) Apparent Radiance Image



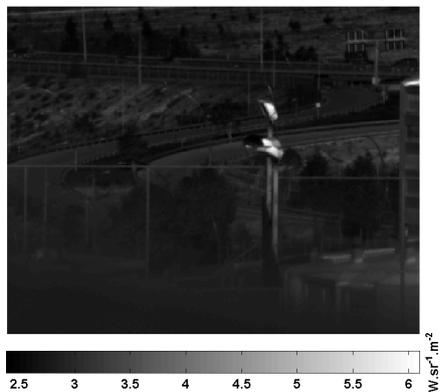
(b) Apparent Radiance Image Distribution



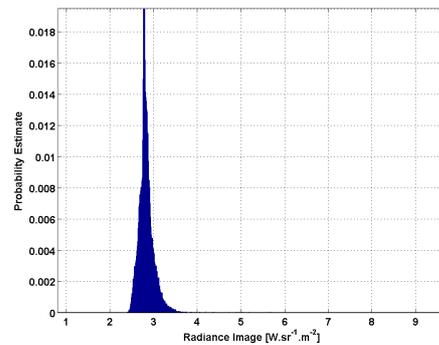
(c) Estimated Source Radiance Image



(d) Estimated Source Radiance Image Distribution



(e) Blackbody Radiance Image



(f) Blackbody Radiance Image Distribution

Figure 8.24: Dual Band Measurement Setup Atmospheric Effects Removal Radiance Images - MWIR Camera - Far Scenario

Moreover, the blackbody radiance image is given in Fig.8.24(e). The distribution of the blackbody image is shown in Fig.8.24(f).

A sample frame of the far scenario for the LWIR camera is analyzed, too. The DN

image captured by the LWIR camera and the result of the temperature calibration for the respective image are shown in Fig.8.25(a) and Fig.8.25(b), respectively. The con-

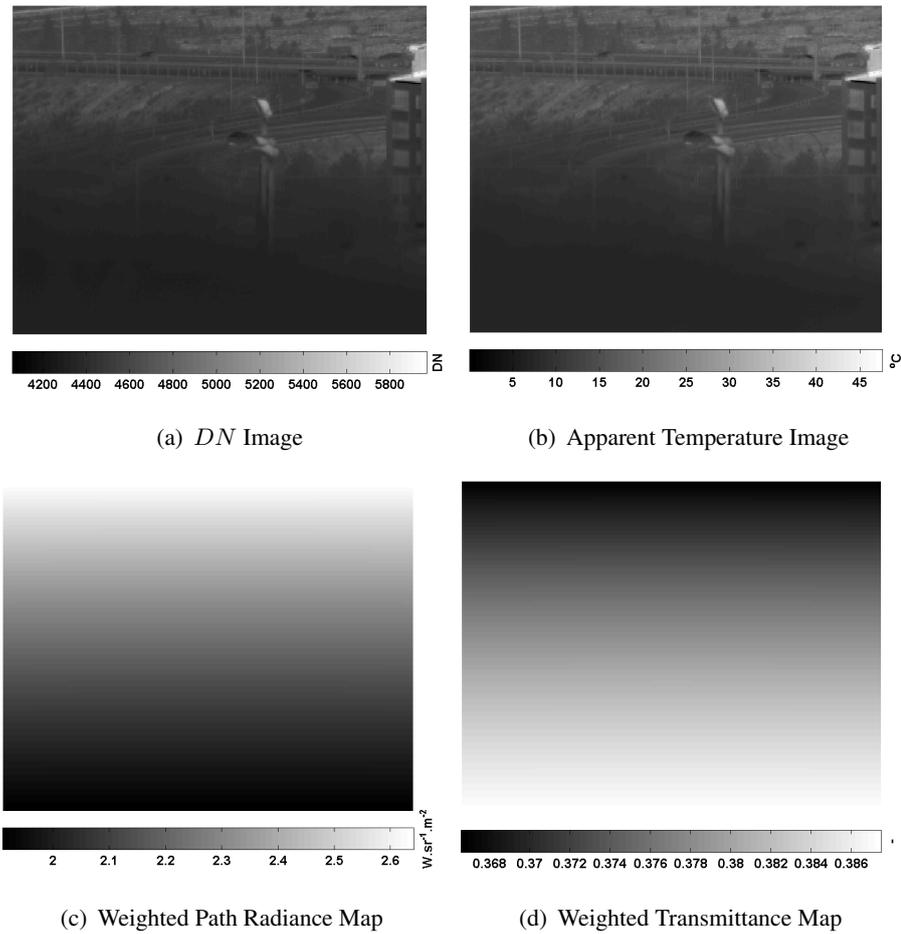
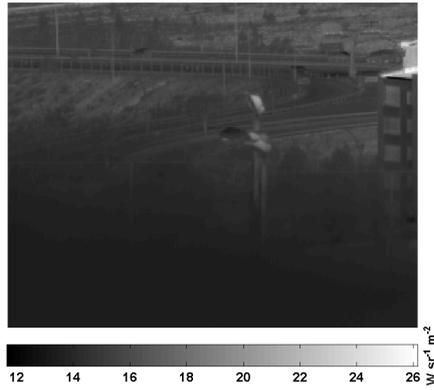


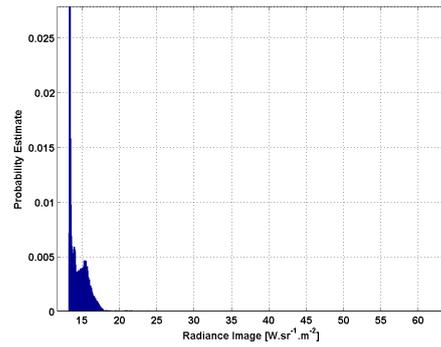
Figure 8.25: Dual Band Measurement Setup Atmospheric Effects Removal Analysis Results - LWIR Camera - Far Scenario

tents of the scene in terms of calibration temperature are in the interval approximately from $0^{\circ}C$ to $47^{\circ}C$. Again, these values are apparent temperatures, not surface temperatures. Therefore, the temperature values of the MWIR and LWIR cameras are not expected to be the same values. The weighted path radiance component for the LWIR camera is shown in Fig.8.25(c). In addition, τ_w for the LWIR camera is shown in Fig.8.25(d).

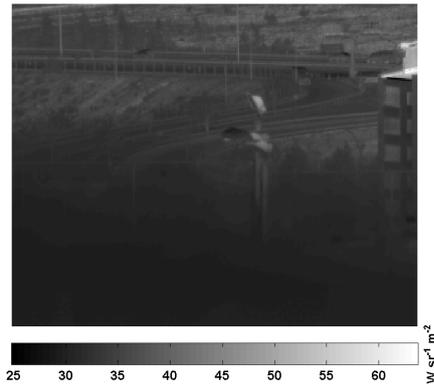
The apparent radiance is shown in Fig.8.26(a). The distribution of the apparent radiance image is shown in Fig.8.26(b). The estimated source radiance and the image distribution are given in Fig.8.26(c) and Fig.8.26(d), respectively. Similar to the MWIR camera case, the distribution of the image is changed due to the removed at-



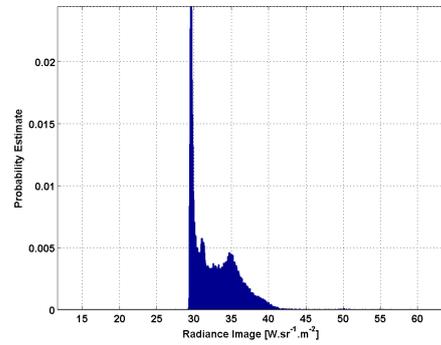
(a) Apparent Radiance Image



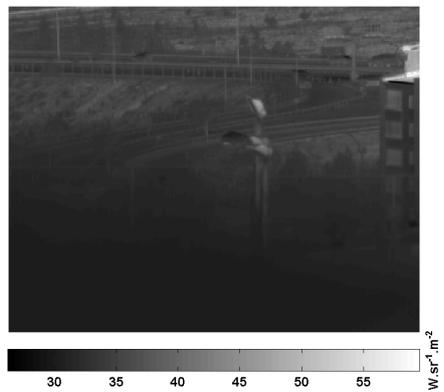
(b) Apparent Radiance Image Distribution



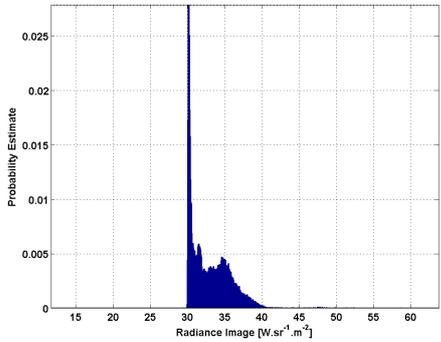
(c) Estimated Source Radiance Image



(d) Estimated Source Radiance Image Distribution



(e) Blackbody Radiance Image



(f) Blackbody Radiance Image Distribution

Figure 8.26: Dual Band Measurement Setup Atmospheric Effects Removal Radiance Images - LWIR Camera - Far Scenario

mospheric effects. Moreover, the blackbody radiance image is given in Fig.8.26(e). The distribution of the blackbody image is shown in Fig.8.26(f).

In addition to the single frame analyses, a global analysis has also been made for all

the recordings of the dual band measurements. Among these measurements first, second and the last measurements in time are the far scenarios, whereas the remaining two measurements are the near scenarios. For the far scenarios, there are 16 recordings. For each recording a ROI, which is 23×8 pixels wide, is specified. The ROI is a region corresponding only to the asphalt in the scene. In these recordings 100 frames are determined such that there is no other object occluding the asphalt. For the near scenarios, there are 11 recordings. For each recording a ROI, which is 23×8 pixels wide, is specified. The ROI is again a region corresponding only to the asphalt in the scene. In these recordings 100 frames are determined such that there is no other object occluding the asphalt. Then the apparent radiance, blackbody radiance and also the estimated source radiance values for each of these pixels in the determined frames are computed similar to the single frame analysis. The temporal and spatial average and also the standard deviation of each scenario for the MWIR and LWIR cameras are computed and shown as an error bar plot in Fig.8.27 and in Fig.8.28, respectively.

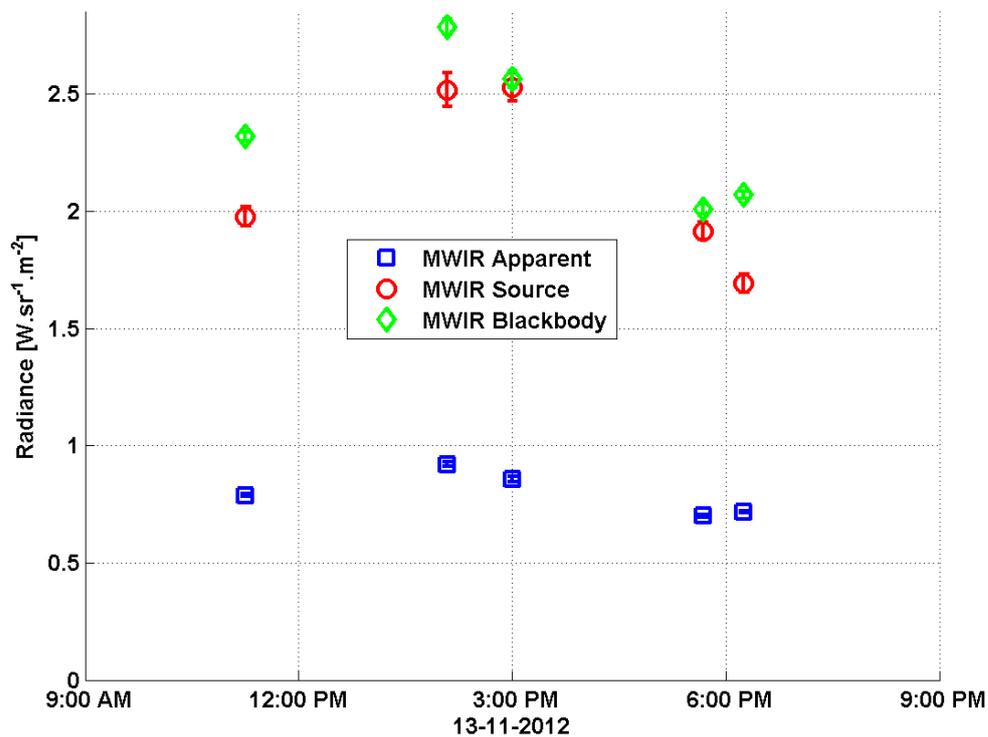


Figure 8.27: Dual Band Measurement Setup Atmospheric Effects Removal Results - MWIR Camera

The apparent radiance values for both cameras are much lower than the blackbody

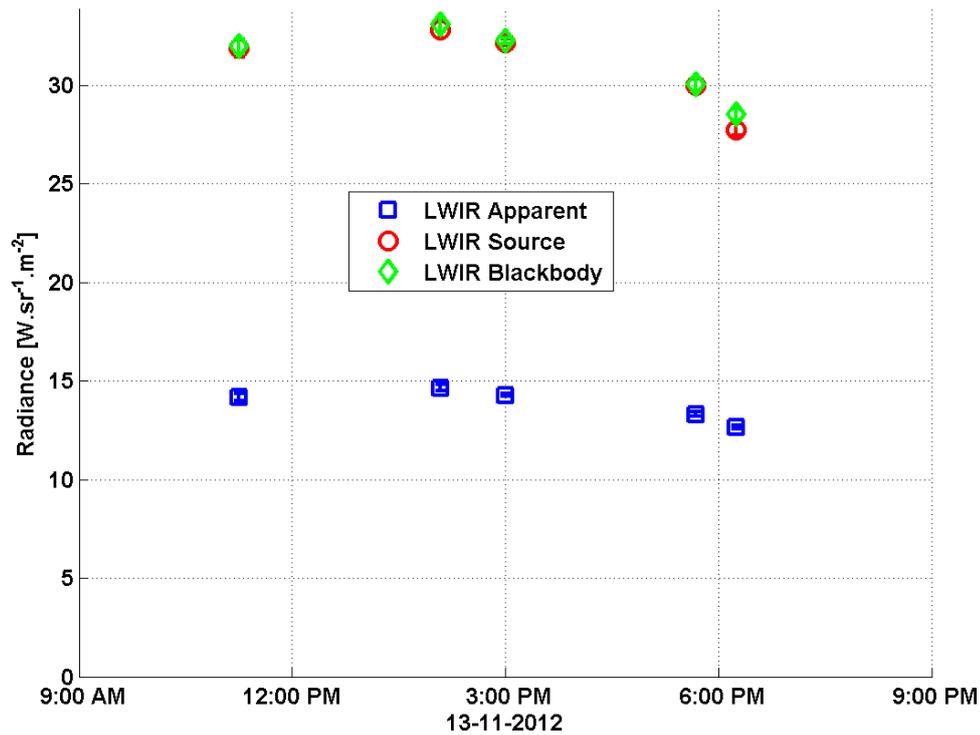
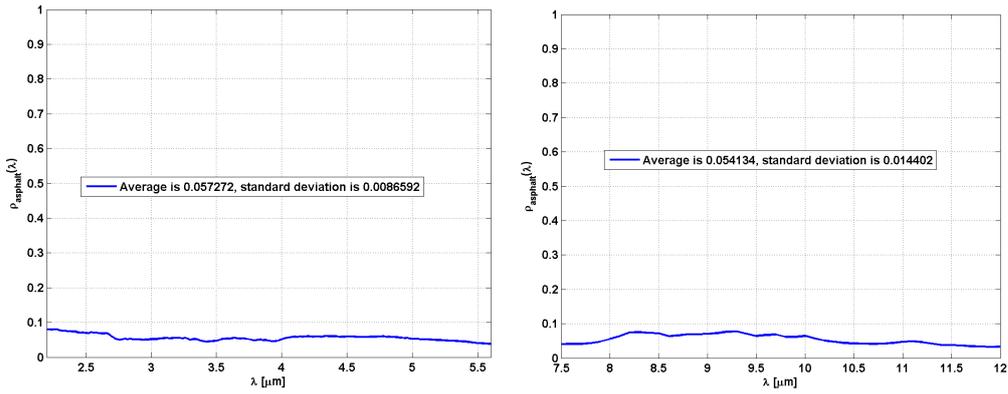


Figure 8.28: Dual Band Measurement Setup Atmospheric Effects Removal Results - LWIR Camera

and estimated source radiance values. This is clearly due to the fact that, the system transfer function of the camera lets approximately half of the incident energy on the lens pass to the detector. As the apparent radiance values contain significant effects of the lens, filter, and the detector; they can be misleading when we consider the scene radiance alone, but not with the sensor. Hence, the apparent radiance values are contained in the graphs just for the illustration of the behavior of the system, but not for the atmospheric effects removal analyses.

ROIs are chosen such that they contain only asphalt because according to spectral library in [135], asphalt behaves almost like a blackbody. The spectral directional hemispherical reflectance of the asphalt in the pass-band of the MWIR and LWIR cameras are given in Fig.8.29(a) and in Fig.8.29(b), respectively. The average directional hemispherical reflectance values in both pass-bands of the cameras are approximately 0.05 and the standard deviation values are around 0.01. Therefore, the reflections from asphalt surface can be neglected. Moreover, the air temperature measurements for the respective times are shown in Fig.8.30. Hence, the dominant difference between the blackbody and the estimated source radiance values are because of



(a) Pass-band of the MWIR Camera

(b) Pass-band of the LWIR Camera

Figure 8.29: Spectral Reflectance of the Asphalt for the Dual Band Measurement Setup

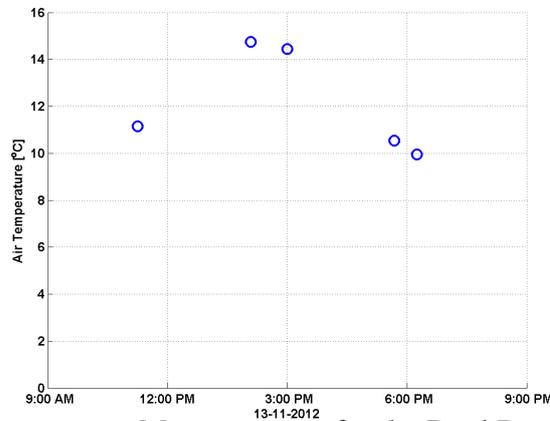


Figure 8.30: Air Temperature Measurements for the Dual Band Measurement Setup

the atmospheric effects. Similar to the single band system case, if there were an ideal atmosphere, the blackbody and the estimated source radiance values are expected to have almost the same values.

In the LWIR band, all the values have the same trend as a function of time and also possess a high correlation with the air temperature measurements. However, in the MWIR band the correlation of the values are lower, especially at the last time slot. During the measurements of the last two time slots, i.e. at 17:40 and 18:15 local times, effectively there isn't any energy coming from neither the sun nor the moon. That means the solar and lunar irradiance values are negligible at these times. As mentioned above, the asphalt behaves almost like a blackbody and so it is a very good absorber and also a radiator. Since there isn't any effective energy source in the scene at the last two time slots, the radiation of the asphalt is not expected to

increase. Moreover, in the steady-state it is expected that the surface temperature of the asphalt and air temperature be close to each other due to the convection. However, according to the blackbody radiance value, the radiation increases at the last time slot with respect to the previous time slot. As the self emission of the asphalt can not increase, the increase can be explained such that it is due to the increase in the path radiance or increase in the atmospheric transmittance. Once the atmospheric effects are removed, it is seen in the estimated source radiance values that the radiation of the asphalt in the last time slot is lower than the previous time slot as expected. As a result, the atmospheric effects removal yields consistent estimates with the diurnal cycle of the extraterrestrial energy both in the MWIR and LWIR bands, whereas the blackbody radiance computations fail to.

8.4 Data Fusion Results

Obviously, data fusion analyses have been made only for the dual band measurement setup. A sample frame of the near scenario is analyzed. The base MWIR image, shown in Fig.8.31(a) and also the registered LWIR image, shown in Fig.8.31(b) are used to obtain the fused image. The fused image is given in Fig.8.31(c). As expected, the fused image contains discriminatory information from both images. For example, the guardhouse on the bottom right corner of the LWIR image cannot be seen in the MWIR image but it is clearly seen in the fused image. Similarly, the light of the car in the bottom left corner of the MWIR image cannot be seen in the LWIR image but it is clearly seen in the fused image. Therefore, the fusion process applied yields an image that has more contrast with respect to the base and registered images.

A sample frame of the far scenario is analyzed, too. The base MWIR image, shown in Fig.8.32(a) and also the registered LWIR image, shown in Fig.8.32(b) are used to obtain the fused image. The fused image is given in Fig.8.32(c). As expected, the fused image contains discriminatory information from both images. For example, the truck on the top right corner of the LWIR image cannot be clearly seen in the MWIR image but it is clearly seen in the fused image. Similarly, the fence in the bottom part of the MWIR image cannot be seen in the LWIR image but it is clearly seen in the fused image. Therefore, the fusion process applied yields an image that has more

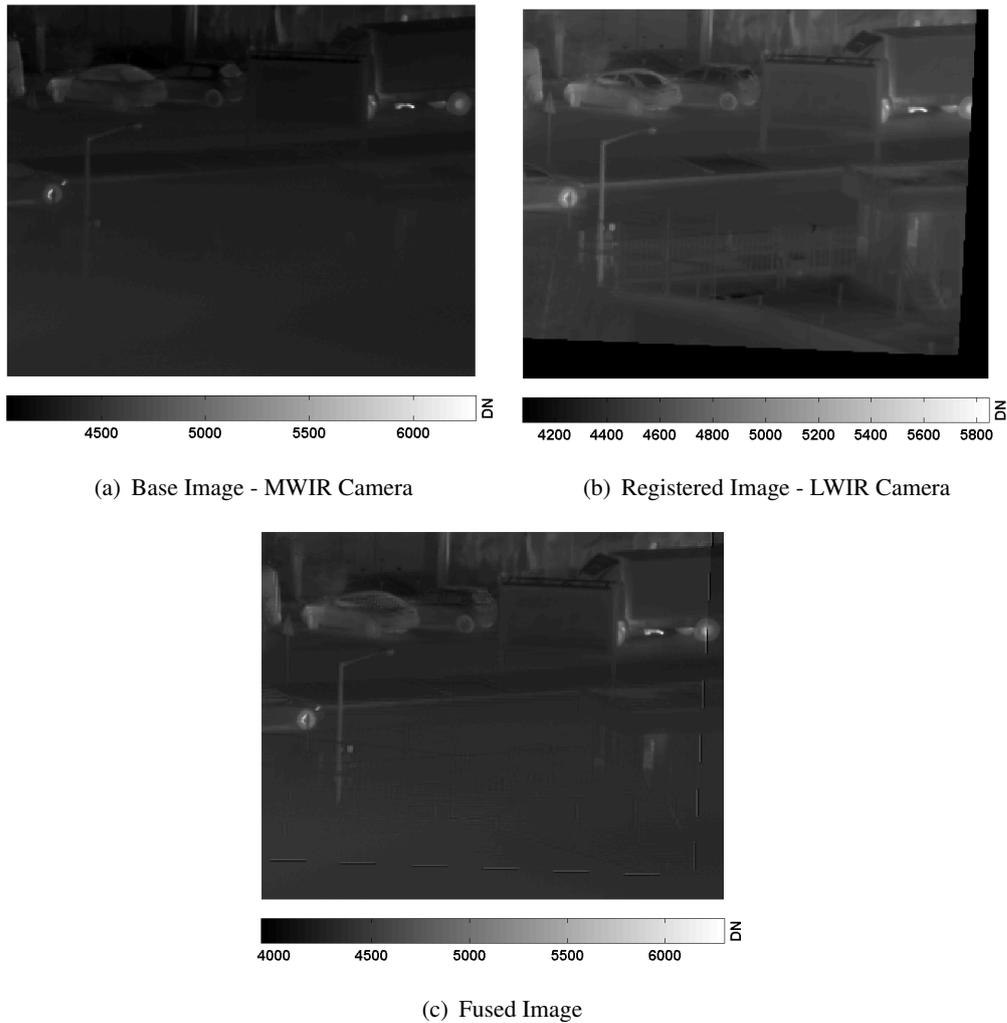
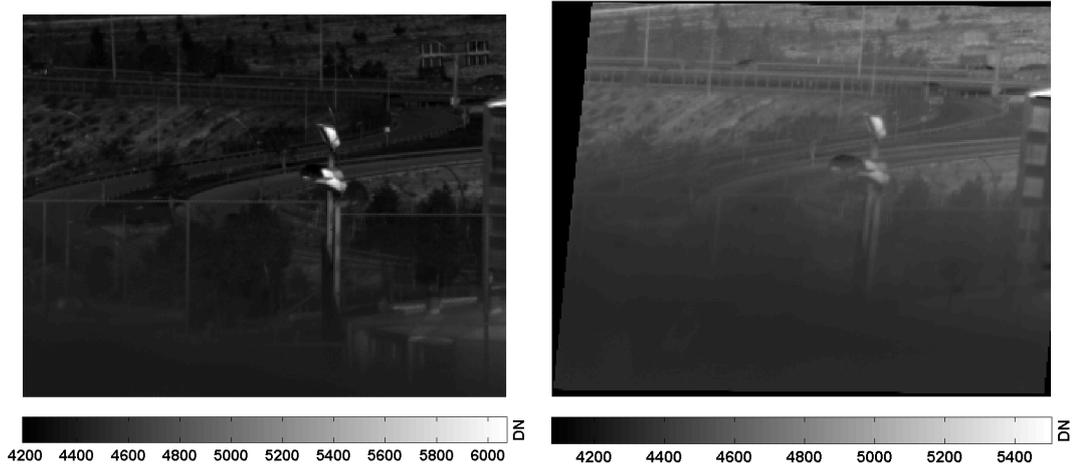


Figure 8.31: Dual Band Measurement Setup Data Fusion Results - Near Scenario

contrast with respect to the base and registered images.

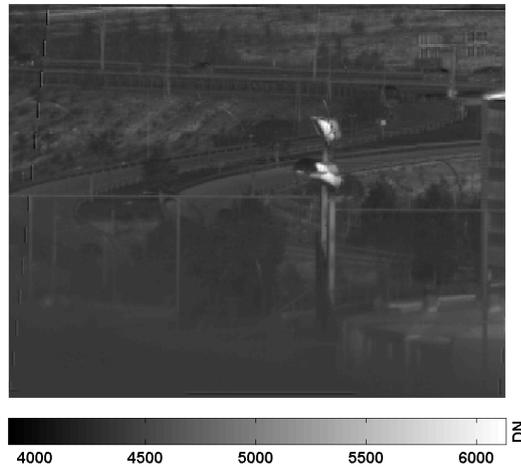
8.5 Moving Vehicle Detection Results

GMM based BG subtraction method is implemented with a maximum number of Gaussian components equal to 10 and the method needs a number of frames in order to adapt the GMM to the scene. Hence, in the subsequent analyses first the training phase of the GMM background subtraction algorithm is completed and then the respective analysis is made.



(a) Base Image - MWIR Camera

(b) Registered Image - LWIR Camera



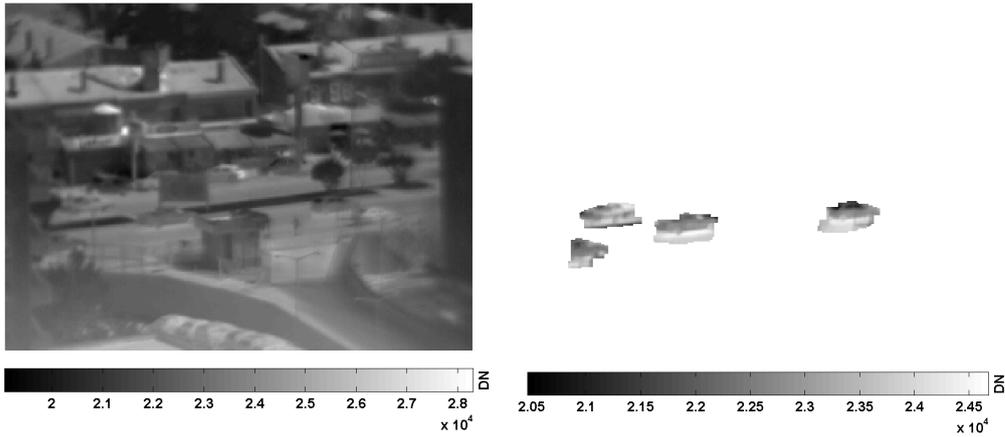
(c) Fused Image

Figure 8.32: Dual Band Measurement Setup Data Fusion Results - Far Scenario

8.5.1 Moving Vehicle Detection Results for the Single Band Measurements

A sample frame of the near scenario is analyzed. The DN image, where there are 4 different moving cars in the scene, is shown in Fig.8.33(a). The segmented image is given in Fig.8.33(b). In the segmented image, all the moving cars are successively separated from the background.

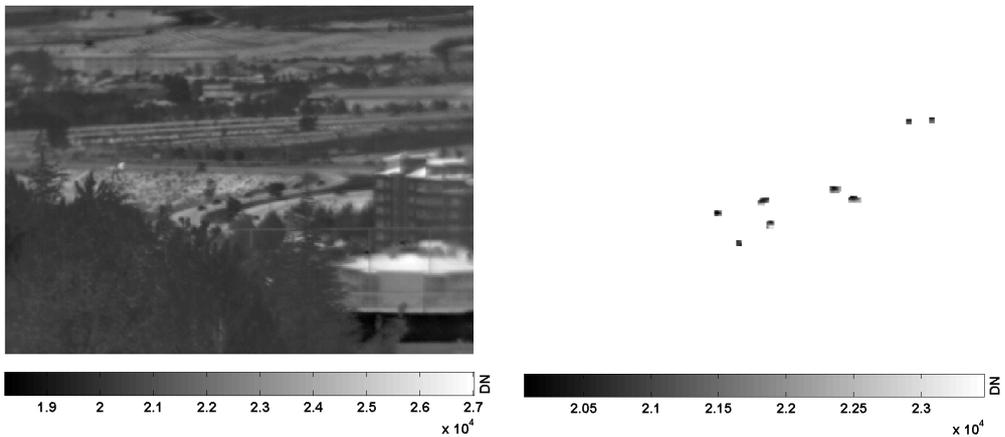
A sample frame of the far scenario is analyzed, too. The DN image is shown in Fig.8.34(a). The segmented image is given in Fig.8.34(b). As the range between the camera and the road is relatively high compared to the near scenario case and the lens on the camera is not big enough, the problem becomes quite difficult even



(a) *DN Image*

(b) Segmented Image

Figure 8.33: Single Band Measurement Setup Moving Vehicle Detection Results - Near Scenario



(a) *DN Image*

(b) Segmented Image

Figure 8.34: Single Band Measurement Setup Moving Vehicle Detection Results - Far Scenario

with a manual segmentation. However, the algorithm successfully separates 8 moving vehicles from the background. The vehicles that could not be separated are the ones, which are relatively far away from the camera and occupy only a few tens of pixels on the captured frame. Therefore, they are treated as noise in the algorithm.

8.5.2 Moving Vehicle Detection Results for the Dual Band Measurements

A sample frame of the near scenario is analyzed. The fused image is shown in Fig.8.35(a), where there are 2 different moving cars in the scene. The segmented

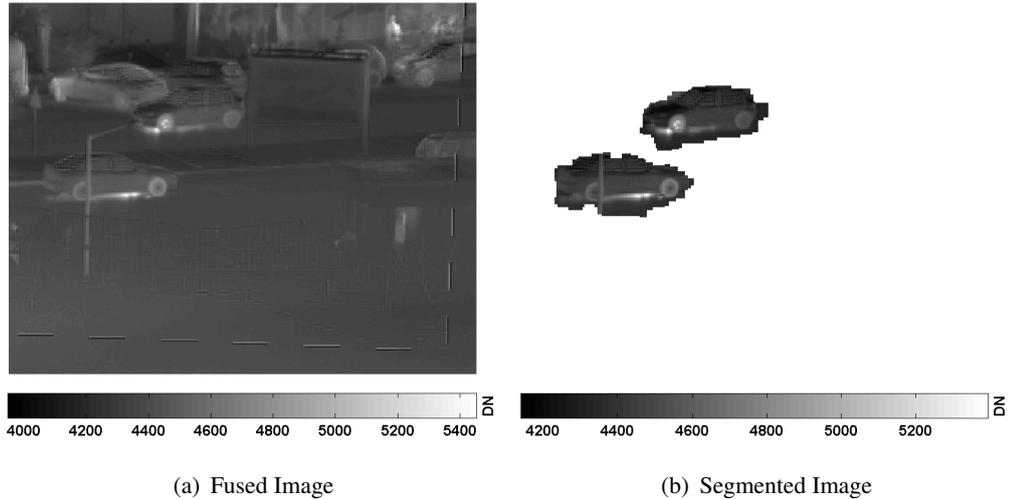


Figure 8.35: Dual Band Measurement Setup Moving Vehicle Detection Results - Near Scenario

image is given in Fig.8.35(b). In the segmented image, all the moving cars are successfully separated from the background.

A sample frame of the far scenario is analyzed, too. The fused image is shown in Fig.8.36(a), where there are 8 different moving vehicles in the scene. The segmented

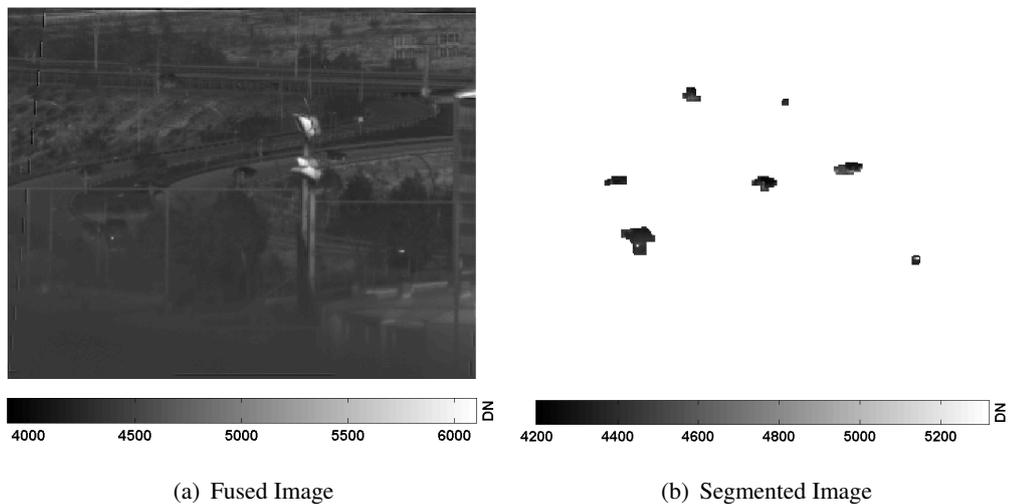


Figure 8.36: Dual Band Measurement Setup Moving Vehicle Detection Results - Far Scenario

image is given in Fig.8.36(b). Although the lens on the camera is bigger than the one in the single band case, the problem is still difficult even with a manual segmentation since the range between the camera and the road is relatively high. However, the algorithm separates 7 moving components from the background. The moving vehicle

that could not be separated is the one, which is relatively far away from the camera and has very similar brightness values with the background. Moreover, due to the increased range, the boundaries of the components are not extracted as good as the boundaries of the near scenario case.

8.6 Classification Results

Four different classes are designated that are car, van, truck/bus, and BG clutter. BG clutter class is used for blobs such as more than one human-being very close to each other that is occluded human and segmentation artifacts due to GMM adaptation to sudden brightness changes like frame drops. Then, all the moving vehicle detection output is labeled manually to one of these classes so as to construct the ground truth class labels. As suggested in [127], obtained feature vectors are normalized to $[0, 1]$. For instance the major axis length of the blob feature vector is normalized among all the feature instances so that the minimum and maximum value of the major axis lengths are 0 and 1, respectively.

Once the ground truth data is obtained and features are normalized, 5-fold cross validation is applied using all the feature vector instances and ground truth data. As a result the optimum regularization and Gaussian RBF kernel parameters are obtained. While searching the optimum parameters, both regularization and RBF kernel parameters took values in the search space 2^i , where $i = -10, -9, \dots, 0, \dots, 9, 10$. Hence, cross validation is evaluated in a 2D grid with $21 \times 21 = 441$ points. The optimum parameter pair is designated as the point with the highest accuracy value.

After finding the optimum SVM parameters, the feature vector instances are randomly separated into two disjoint sets for training and test. While separating the feature vectors, occluded vehicle feature vectors are not included in the training set but they are used in the test set. Hence, the failure of the moving vehicle detection step to discriminate a vehicle from another vehicle or background is taken into account in the performance evaluations, which is the case in practice. Then, the feature vector instances and the respective ground truth data of the training set is used to train an SVM model with the optimum parameters of the cross validation step. Afterwards,

SVM model is tested with the test data. To increase the accuracy of the statistical distribution, 100 Monte Carlo runs are executed for the data separation, training and test steps. At the end of Monte Carlo runs, all the results are merged as a confusion matrix. Then, accuracy metric, precision metric for each class, and recall metric for each class are computed using the confusion matrix. Clearly, while implementing the above methodology for the reference and the proposed systems, both systems are evaluated using the same randomly selected data set. The only difference of the two systems is that they use different features and as a result they have different optimum parameters.

In all the measurements, the representative data of each class are unbalanced. Number of cars are much larger than the total number of the rest of the classes. Therefore, using the strategy in [127], different regularization parameters are assigned to each class both in the cross validation and training steps. The reported optimum regularization parameters in the following subsections correspond to the parameter of the class with the largest number of instances. Hence, the parameters of the remaining classes are obtained by weighting this regularization parameter inversely proportional to the number of instances. That is the regularization parameter of class i , C_i , is computed using the following equalities

$$C_i = w_i C \quad (8.2)$$

where the weight factor of class i , w_i , is calculated with

$$w_i = \left[\frac{\max_i N_i}{N_i} \right] \quad (8.3)$$

where N_i is the number of samples of class i and $[.]$ is round-to-integer operator.

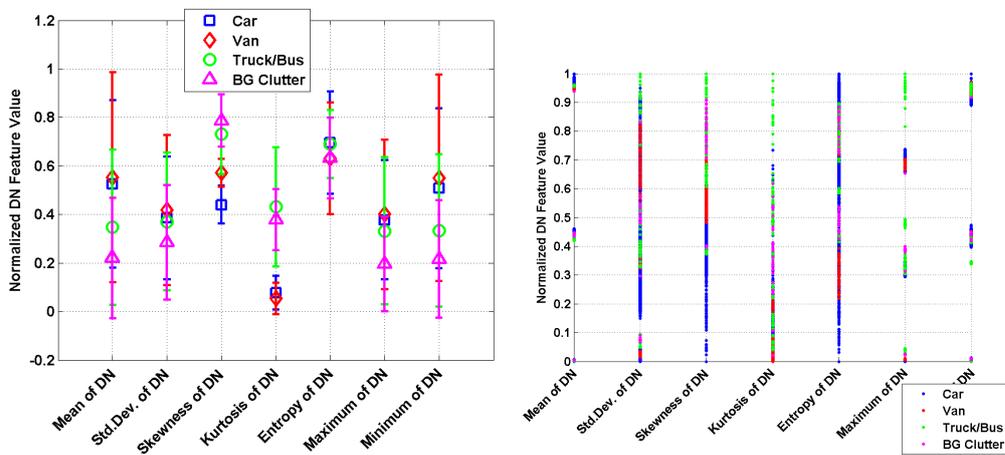
In the following subsections, to emphasize the need for a radiometric calibration, classification with source radiance estimate features and classification with DN features are given first. Then, to motivate the use of atmospherically corrected data rather than uncorrected data, classification with source radiance estimate features and classification with blackbody radiance features are compared. In the following two subsections, the performance of the proposed system is compared with the reference system. In the reference system only blob features are used, whereas both blob and radiometric features are used in the proposed system. The results obtained with the

single band measurement system and also the results obtained with the dual band measurement system are given. Finally, to clarify the effect of moving vehicle detection performance on classification, all the experiments are repeated without using any occluded samples in the tests.

8.6.1 Effects of Radiometric Calibration on Vehicle Classification

In order to assess the effects of radiometric calibration on vehicle classification, two systems for the single band measurement system are used. In the first system, the features are obtained from the *DN* frames. Similar to the radiometric features mean, standard deviation, skewness, kurtosis, entropy, minimum, and maximum value of the *DN* frames are extracted as the features of the first system and called *DN* features. In the second system, only the radiometric features are used.

The average and the standard deviation of the normalized *DN* feature instances are calculated. Statistics of each component of the *DN* feature vectors are given as error bar plot in Fig.8.37(a) for the single band measurement setup. All instances of the normalized *DN* features are also shown in Fig.8.37(b). Similarly, the statistics and also

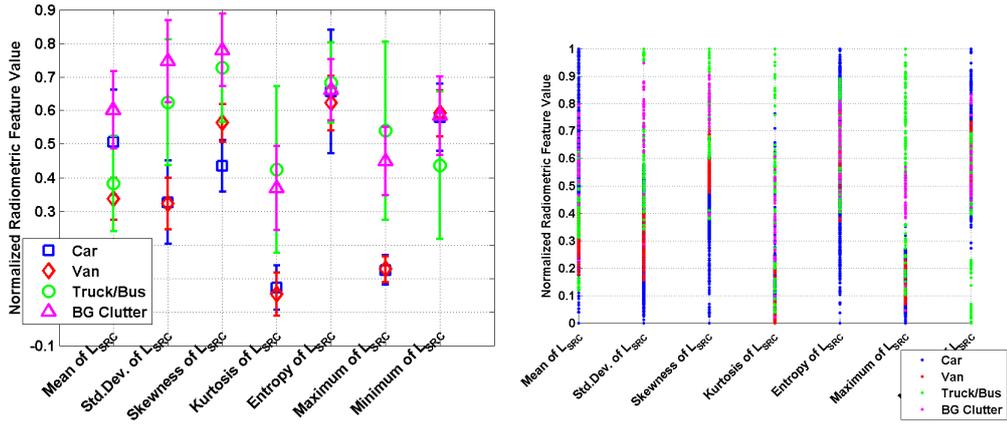


(a) Normalized *DN* Features Statistics

(b) Normalized *DN* Features

Figure 8.37: Single Band Measurement Setup *DN* Feature Extraction Results

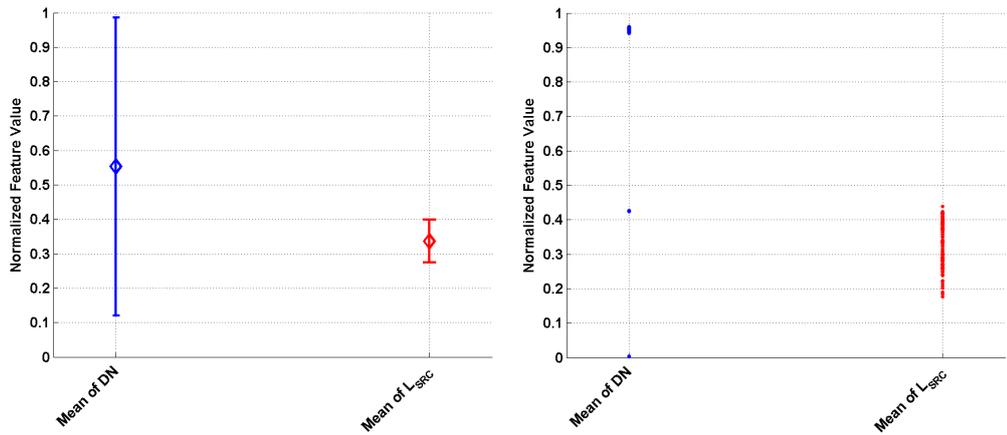
all the instances of the radiometric features are shown in Fig.8.38(a) and Fig.8.38(b), respectively. Furthermore, in order to have a closer look at the distribution of the features, a sample member that is the mean component of the *DN* and radiometric features are chosen for the van class. The average and the standard deviation of the



(a) Normalized Radiometric Features Statistics (b) Normalized Radiometric Features

Figure 8.38: Single Band Measurement Setup Radiometric Feature Extraction Results

normalized mean of DN image and also the normalized mean of source radiance estimate image instances are calculated. Statistics of the respective normalized feature components are given as an error bar plot in Fig.8.39(a). Moreover, all instances of the respective normalized feature components are also shown in Fig.8.39(b). Since

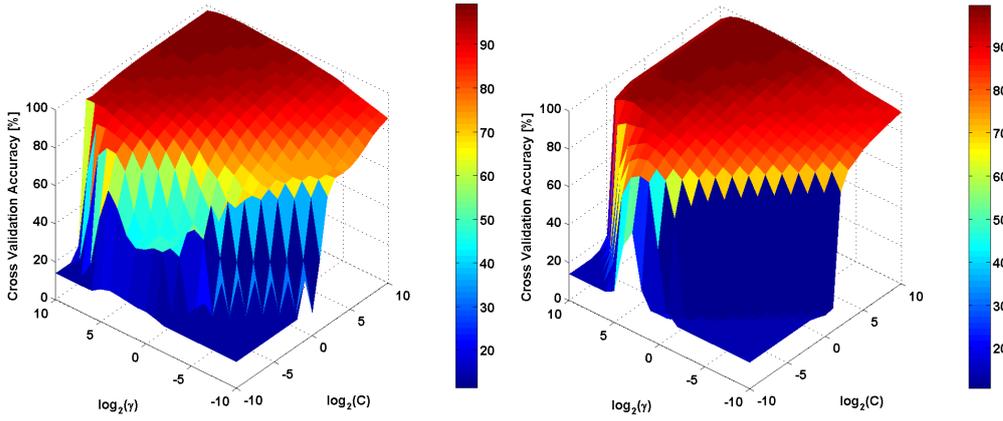


(a) Normalized Features Statistics for Van (b) Normalized Features for Van

Figure 8.39: Single Band Measurement Setup DN vs Radiometric Feature Comparison for the Van Class

the measurements in this setup are taken using three different IT values, the variance of the DN features are larger than the variance of the radiometric features. As a result, a performance gain is expected with the radiometric features.

For the first system, the optimum regularization parameter is found to be $C = 256$ and the optimum Gaussian RBF kernel parameter is found to be $\gamma = 128$, where the related search space and accuracies are shown as a surface plot in Fig.8.40(a). The



(a) First System

(b) Second System

Figure 8.40: Classification with DN vs. Radiometric Features Cross Validation Results

confusion matrix is given in Table 8.1. The precision and recall metrics obtained for each class and also the accuracy of the system are given in Table 8.2 together with the standard deviation values among Monte Carlo runs.

Table 8.1: Classification with DN Features Confusion Matrix

		Prediction			
		<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Real	<i>Car</i>	221150	4527	165	250
	<i>Van</i>	1278	18251	0	170
	<i>Truck/Bus</i>	173	1	1683	9
	<i>BG clutter</i>	16418	5584	2153	28321

Table 8.2: Classification with DN Features Performance Metrics

	<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Precision	92.52%	64.35%	42.06%	98.51%
σ_{prec}	0.78%	3.98%	28.52%	29.02%
Recall	97.81%	92.65%	90.19%	53.97%
σ_{rec}	0.26%	2.43%	17.74%	20.84%
Accuracy	89.76%			
σ_{acc}	0.4553%			

For the second system, the optimum regularization parameter is found to be $C = 64$ and the optimum Gaussian RBF kernel parameter is found to be $\gamma = 32$, where the related search space and accuracies are shown in Fig.8.40(b). The confusion matrix is given in Table 8.3. The precision and recall metrics obtained for each class and also the accuracy of the system are given in Table 8.4 together with the standard deviation values among Monte Carlo runs.

Table 8.3: Classification with Radiometric Features Confusion Matrix

		Prediction			
		<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Real	<i>Car</i>	219439	4860	169	1624
	<i>Van</i>	1333	18323	0	43
	<i>Truck/Bus</i>	137	5	1723	1
	<i>BG clutter</i>	12793	2231	195	37257

Table 8.4: Classification with Radiometric Features Performance Metrics

	<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Precision	93.9%	72.08%	82.56%	95.71%
σ_{prec}	0.48%	4.75%	7.85%	9.57%
Recall	97.06%	93.01%	92.34%	71%
σ_{rec}	0.45%	1.83%	10.21%	12.85%
Accuracy	92.21%			
σ_{acc}	0.53%			

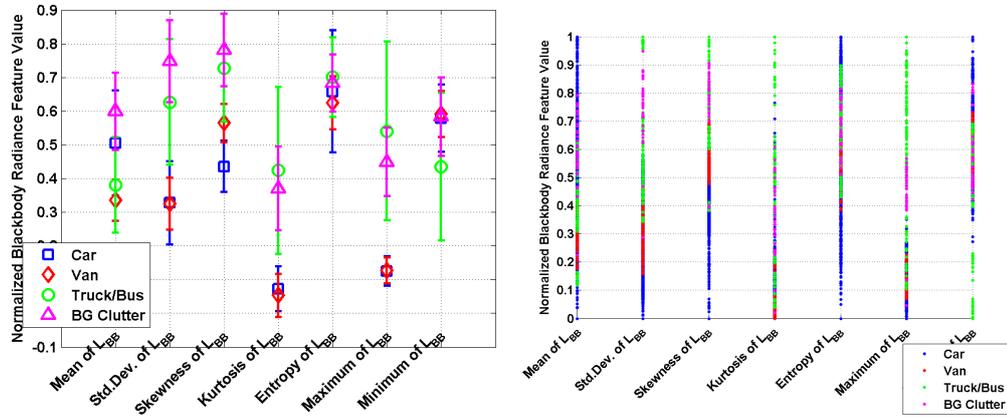
The use of radiometric calibration in classification clearly improves the overall classification accuracy with respect to the case with raw data, i.e. DN features. The second system is more successful in terms of recall in the classification of all the classes except car class. In terms of precision, the second system is better for all the classes. When the entries of the confusion matrices are investigated, it is seen that the correct predictions of the second system is more for all the classes except the car class. However, the decrease of correct predictions of car class is negligible when the increase of the other classes are considered.

A naive look to the results would be the use of only DN features are quite good. However, in the newly emerging IR cameras the IT value used during the measurements are started to be dynamically adapted to the scene. That is, the IT value for each frame is adjusted depending on the scene. For instance, if a very hot entity, e.g. the plume of a truck, enters the FOV of the camera, the IT value of the camera is decreased so that no saturation from upper DN limit occurs. Similarly, if a very cold entity, e.g. a car with snow on top of it, enters the FOV, the IT value is increased so that no saturation from lower DN limit occurs. Therefore, in such a configuration the value of DN changes from frame to frame and so DN values cannot be trusted in classification. However, once all these DN values are mapped to a physical quantity like radiance with the help of radiometric calibration, the consistency of the value for the same object within frames is obtained. Thus these radiometric quantities can be used in the classification. Similarly, the discussion above is valid for the case of IR cameras that change the neutral density (energy limiting) filters adaptively during the measurements. When a very hot entity enters the FOV, a neutral density filter with a relatively lower transmittance characteristic is dynamically put on the camera. In this situation, as the energy is decreased to prevent saturation from upper DN limits, the DN values of the scene decreases as is the case with IT adaptation. However, the effect of dynamic filter adjustment on the frame data can be drastically decreased by mapping the DN values to radiance values with the help of radiometric calibration. To conclude, the proposed method has the potential to be used for vehicle classification with IR cameras that has the IT adaptation and/or neutral density filter adaptation capabilities.

8.6.2 Effects of Atmospheric Correction on Vehicle Classification

In order to evaluate the effects of the proposed atmospheric correction method on vehicle classification, two systems for the single band measurement system are used. In the first system, the features are obtained from the blackbody radiance frames. Similar to the radiometric features mean, standard deviation, skewness, kurtosis, entropy, minimum, and maximum value of the blackbody radiance frames are extracted as the feature vector of the first system and called blackbody radiance features. In the second system, only the radiometric features are used.

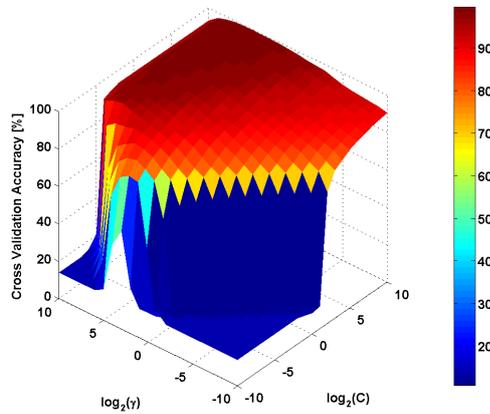
The average and the standard deviation of the normalized blackbody feature instances are calculated. Statistics of each component of the blackbody feature vectors are given as error bar plot in Fig.8.41(a) for the single band measurement setup. All instances of the normalized blackbody features are also shown in Fig.8.41(b).



(a) Normalized Blackbody Features Statistics (b) Normalized Blackbody Features

Figure 8.41: Single Band Measurement Setup Blackbody Feature Extraction Results

For the first system, the optimum regularization parameter is found to be $C = 16$ and the optimum Gaussian RBF kernel parameter is found to be $\gamma = 64$, where the related search space and accuracies are shown as a surface plot in Fig.8.42(a). The



(a) First System

Figure 8.42: Classification with Blackbody Radiance vs. Radiometric Features Cross Validation Results

confusion matrix is given in Table 8.5. The precision and recall metrics obtained for each class and also the accuracy of the first system are given in Table 8.6 together with the standard deviation values among Monte Carlo runs.

Table 8.5: Classification with Blackbody Radiance Features Confusion Matrix

		Prediction			
		Car	Van	Truck/Bus	BG clutter
Real	Car	222167	3785	117	23
	Van	1478	18213	0	8
	Truck/Bus	287	0	1579	0
	BG clutter	23470	1177	0	27829

Table 8.6: Classification with Blackbody Radiance Features Performance Metrics

	Car	Van	Truck/Bus	BG clutter
Precision	89.8%	78.59%	93.1%	99.89%
σ_{prec}	0.57%	3.09%	6.3%	7.1%
Recall	98.26%	92.46%	84.62%	53.03%
σ_{rec}	0.29%	1.62%	16.06%	18.24%
Accuracy	89.89%			
σ_{acc}	0.53%			

For the second system, the results are the same with the ones in Sec.8.6.1. Hence, the optimum regularization parameter is found to be $C = 64$ and the optimum Gaussian RBF kernel parameter is found to be $\gamma = 32$, where the related search space and accuracies have already been shown in Fig.8.40(b). The confusion matrix has been given in Table 8.3. The precision and recall metrics obtained for each class and also the accuracy of the second system have already been given in Table 8.4 together with the standard deviation values among Monte Carlo runs.

The use of atmospherically corrected radiance data in classification clearly improves the overall classification accuracy with respect to the case with uncorrected radiance data, i.e. blackbody radiance features. The second system is more successful in terms of recall in the classification of all the classes except car class. The situation with the precision values is the exact opposite of the situation with the recall values. When the entries of the confusion matrix corresponding to the BG clutter class are considered, it is seen that the second system is much more successful in terms of

correct prediction. This is explained in a way that the atmospheric effects distribution on the frame cannot be neglected. Since the samples of BG clutter class may have blobs that spread in a large area on the frame, the distribution of the atmospheric effects on the blob becomes important. Hence, the number of correct predictions of the samples of BG clutter class are greater with the second system. In other words, if no atmospheric correction is applied, the distribution of the blackbody radiance are less discriminative than the distribution of the atmospherically corrected radiance data, i.e. source radiance estimate. Similar arguments apply to the van and truck/bus classes since the second system performs better in terms of correct prediction of the respective classes. The increase of the correct predictions for these three classes come with a decrease in the correct prediction of car class. However, the decrease is negligible when the increase of the other classes are considered.

Although a performance gain is obtained with the proposed atmospheric correction approach, it shall be noted that only the blackbody radiance values even yield high overall accuracy. To the current knowledge of the author, this thesis is the first that uses the blackbody radiance values in vehicle classification like it is also the first that uses the source radiance estimate values in vehicle classification. In terms of computational complexity, the difference between the blackbody radiance computation and source radiance estimation is a single subtraction and a division for each pixel in the blob. Clearly, the subtracted value is the weighted path radiance component that can be precomputed and stored in the memory. Similarly, the divisor is the weighted transmittance value and again transmittance values can be precomputed. As a result, there is a tradeoff between the overall accuracy performance and computational complexity for the proposed systems. However, if the computational complexity is not critical, the atmospheric correction may be favored since it possesses higher accuracy and also if there is any additional step following the classification, atmospherically corrected radiance data has the high potential to become very valuable for the sequential tasks.

8.6.3 Classification Results for the Single Band Measurements

For the single band measurements, the near scenario is executed according to the aforementioned classification procedure. The reference system that is shown in Fig.1.1 is used. In the reference system, only blob features are used. In the proposed system, both blob and radiometric features are used.

The average and the standard deviation of the normalized blob feature instances are calculated. Statistics of each component of the blob feature vectors are given as error bar plot in Fig.8.43(a) for the single band measurement setup. All instances of the normalized blob features are also shown in Fig.8.43(b).

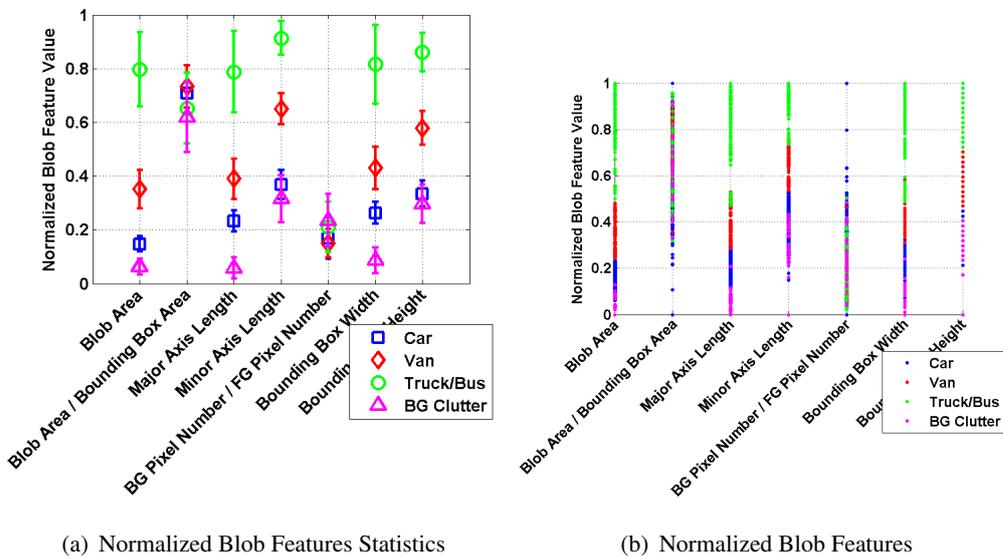
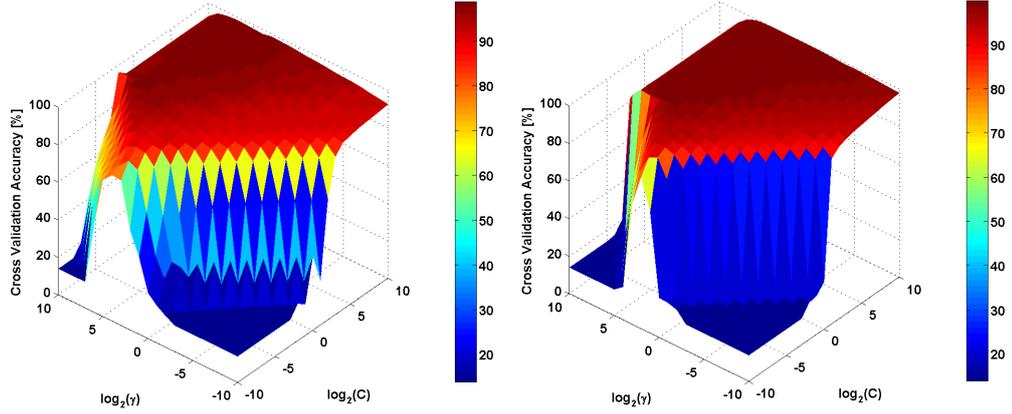


Figure 8.43: Single Band Measurement Setup Blob Feature Extraction Results

The optimum regularization parameter of the reference system is found to be $C = 16$ and the optimum Gaussian RBF kernel parameter is found to be $\gamma = 32$, where the related search space and accuracies are shown as a surface plot in Fig.8.44(a). The confusion matrix is given in Table 8.7. The precision and recall metrics obtained for each class and also the accuracy of the reference system are given in Table 8.8 together with the standard deviation values among Monte Carlo runs.

The proposed system that is shown in Fig.1.2 is used. The optimum regularization parameter is found to be $C = 1024$ and the optimum Gaussian RBF kernel parameter is found to be $\gamma = 16$, where the related search space and accuracies are shown in Fig.8.44(b). The confusion matrix is given in Table 8.9. The precision and recall



(a) Reference Method (b) Proposed Method
 Figure 8.44: Single Band Measurement Setup Cross Validation Results

Table 8.7: Single Band Measurement Setup Reference Method Confusion Matrix

		Prediction			
		<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Real	<i>Car</i>	202573	16449	414	6656
	<i>Van</i>	913	18726	0	60
	<i>Truck/Bus</i>	394	0	1468	4
	<i>BG clutter</i>	9855	1729	0	40892

Table 8.8: Single Band Measurement Setup Reference Method Performance Metrics

	<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Precision	94.78%	50.74%	78%	85.89%
σ_{prec}	2.14%	9.94%	13.14%	9.22%
Recall	89.6%	95.06%	78.67%	77.93%
σ_{rec}	5.27%	3.9%	9.67%	9.1%
Accuracy	87.85%			
σ_{acc}	2.88%			

metrics obtained for each class and also the accuracy of the proposed system are given in Table 8.10 together with the standard deviation values among Monte Carlo runs.

Table 8.9: Single Band Measurement Setup Proposed Method Confusion Matrix

		Prediction			
		<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Real	<i>Car</i>	222420	3601	71	0
	<i>Van</i>	1103	18596	0	0
	<i>Truck/Bus</i>	185	0	1681	0
	<i>BG clutter</i>	21528	1160	0	29788

Table 8.10: Single Band Measurement Setup Proposed Method Performance Metrics

	<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Precision	90.7%	79.62%	95.95%	100%
σ_{prec}	0.76%	4.53%	4.26%	5.28%
Recall	98.38%	94.4%	90.09%	56.76%
σ_{rec}	0.59%	1.65%	16.16%	18.33%
Accuracy	90.79%			
σ_{acc}	0.75%			

The proposed system clearly improves the overall classification accuracy with respect to the reference system. The proposed system is more successful in terms of recall in the classification of the car and truck/bus classes. On the other hand, it performs almost the same for the van class and performs worse for the BG clutter class. This is because, BG clutter class contains a broad group of objects such as occluded human, segmentation artifacts, noise, etc. As a result, the distribution of these objects show a broad range in terms of radiance as their characteristics are much more different. Therefore, the inclusion of the radiometric features decreases the classification performance for the BG clutter. However, in practice the most critical classes are the remaining three classes and the proposed method is shown to be superior for these classes. This is explained as, these specific classes have similar radiance distribution within the same class and so the radiometric features help to increase the performance.

Some examples of the misclassifications of the reference method are given in Fig.8.45. In Fig.8.45(a) and Fig.8.45(b), the car is occluded with a van in the scene. Hence, the

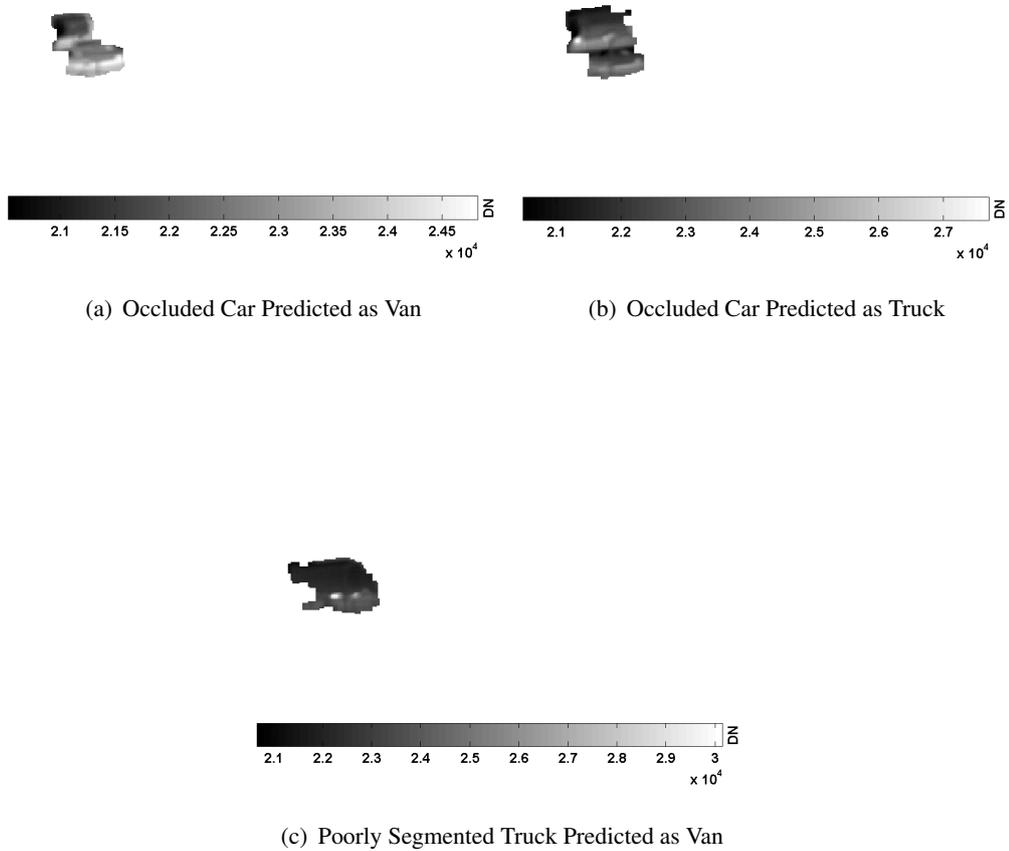


Figure 8.45: Single Band Measurement Setup Misclassification Examples

system that uses only blob features classifies these blobs as van and truck, respectively. This is simply because the occlusion makes the blob appear larger than the blob of a non-occluded car. In Fig.8.45(c), rear part of the truck is designated as background in the vehicle detection step. Hence, the truck appears smaller than its real size. As a result, the reference method classifies the blob of the truck as a van. On the other hand, the proposed method that uses blob features together with the radiometric features classifies all the three examples correctly. This fact highlights the accuracy gain obtained with the use of radiometric features on vehicle classification.

8.6.4 Classification Results for the Dual Band Measurements

For the dual band measurements, the near scenario at local time 15:00 is executed according to the aforementioned classification procedure. The reference system that is shown in Fig.1.3 is used. In the reference system, only blob features are used. In the proposed system, both blob and radiometric features are used.

The average and the standard deviation of the normalized blob feature instances are calculated. Statistics of each component of the blob feature vectors are given as error bar plot in Fig.8.46(a) for the dual band measurement setup. All instances of the normalized blob features are also shown in Fig.8.46(b). Similarly, the statistics

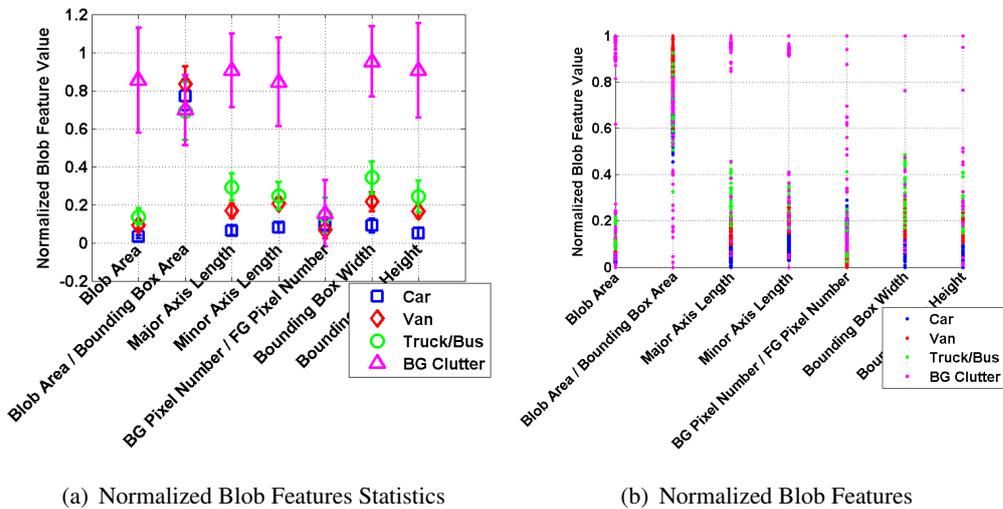
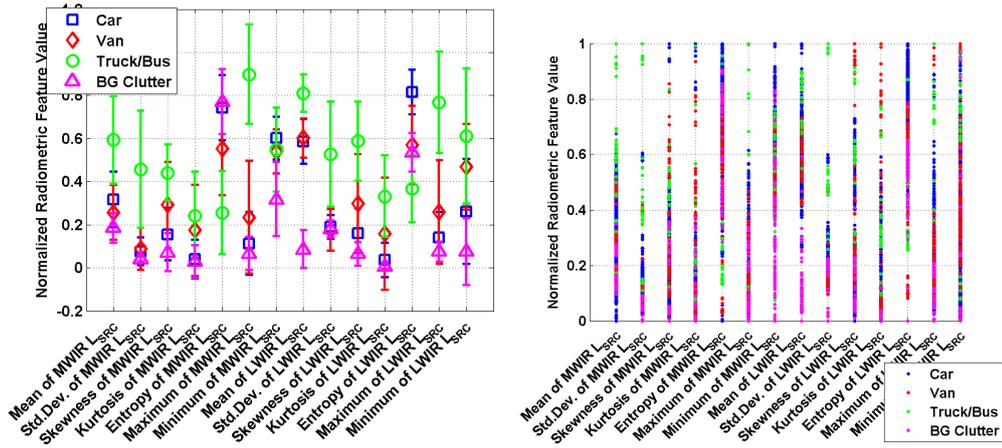


Figure 8.46: Dual Band Measurement Setup Blob Feature Extraction Results

and also all the instances of the radiometric features are shown in Fig.8.47(a) and Fig.8.47(b), respectively.

The optimum regularization parameter of the reference system is found to be $C = 128$ and the optimum Gaussian RBF kernel parameter is found to be $\gamma = 64$, where the related search space and accuracies are shown as a surface plot in Fig.8.48(a). The confusion matrix is given in Table 8.11. The precision and recall metrics obtained for each class and also the accuracy of the reference system are given in Table 8.12 together with the standard deviation values among Monte Carlo runs.

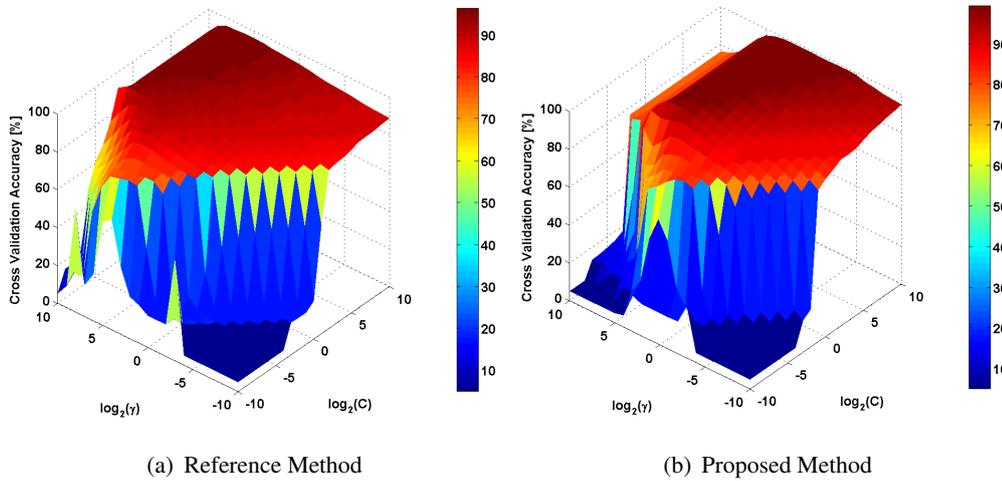
The proposed system that is shown in Fig.1.4 is used. The optimum regularization parameter is found to be $C = 16$ and the optimum Gaussian RBF kernel parameter



(a) Normalized Radiometric Features Statistics

(b) Normalized Radiometric Features

Figure 8.47: Dual Band Measurement Setup Radiometric Feature Extraction Results



(a) Reference Method

(b) Proposed Method

Figure 8.48: Dual Band Measurement Setup Cross Validation Results

Table 8.11: Dual Band Measurement Setup Reference Method Confusion Matrix

		Prediction			
		<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Real	<i>Car</i>	103100	8351	7955	9071
	<i>Van</i>	3906	14544	2636	4485
	<i>Truck/Bus</i>	430	0	6278	57
	<i>BG clutter</i>	8515	491	5964	9895

is found to be $\gamma = 2$, where the related search space and accuracies are shown in Fig.8.48(b). The confusion matrix is given in Table 8.13. The precision and recall metrics obtained for each class and also the accuracy of the proposed system are

Table 8.12: Dual Band Measurement Setup Reference Method Performance Metrics

	<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Precision	88.92%	62.19%	27.5%	42.09%
σ_{prec}	7.08%	7.66%	28.77%	4.63%
Recall	80.25%	56.88%	92.8%	39.79%
σ_{rec}	8.05%	4.86%	4.94%	13.77%
Accuracy	72.07%			
σ_{acc}	5.05%			

given in Table 8.14 together with the standard deviation values among Monte Carlo runs.

Table 8.13: Dual Band Measurement Setup Proposed Method Confusion Matrix

		Prediction			
		<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Real	<i>Car</i>	112185	10216	3974	2102
	<i>Van</i>	2891	18920	1218	2542
	<i>Truck/Bus</i>	104	45	6616	0
	<i>BG clutter</i>	2068	1883	743	20171

Table 8.14: Dual Band Measurement Setup Proposed Method Performance Metrics

	<i>Car</i>	<i>Van</i>	<i>Truck/Bus</i>	<i>BG clutter</i>
Precision	95.68%	60.91%	52.71%	81.29%
σ_{prec}	2.14%	6.36%	19.49%	4.34%
Recall	87.32%	73.99%	97.8%	81.12%
σ_{rec}	2.6%	7.22%	1.95%	12.76%
Accuracy	85.04%			
σ_{acc}	1.98%			

The proposed system improves the overall classification accuracy substantially with

respect to the reference system. The proposed system is more successful in terms of recall in the classification of all the classes. Moreover, the precision is better for all the classes except the van. However, the decrease in the precision for the van class is negligible when compared to the achievements for the rest of precision and recall values. In addition, the accuracy improvement is much more compared to the single band case as expected. Clearly, the dual band measurement system is more complex as it requires one more camera and additional registration and fusion steps in the algorithm pipeline. Therefore, it is concluded that when selecting between the single band and dual band proposed systems, there is a tradeoff between the performance and both computational and hardware complexities.

Some examples of the misclassifications of the reference method are given in Fig.8.49. In Fig.8.49(a), the car is occluded with another car in the scene and so the blob appears

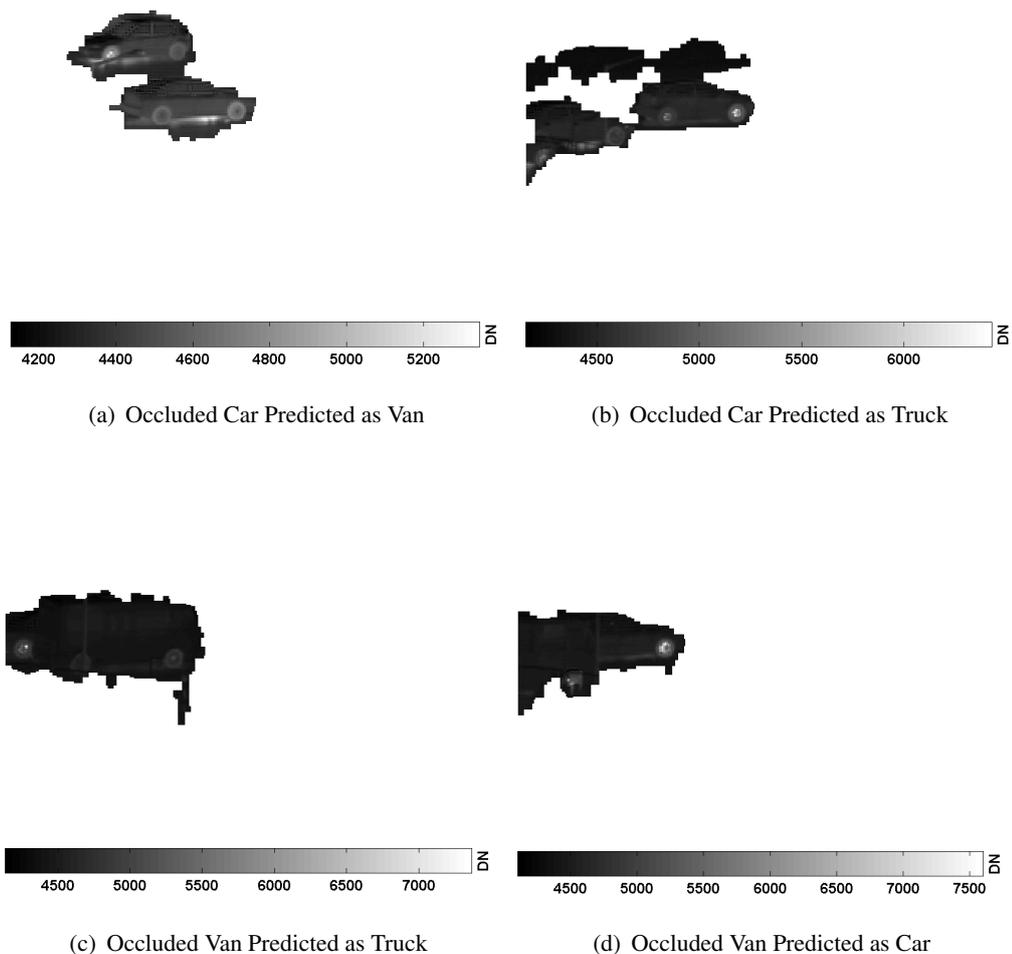


Figure 8.49: Dual Band Measurement Setup Misclassification Examples

larger than the blob of a non-occluded car. Hence, the reference method misclassifies the blob as if it were a van. In Fig.8.49(b), the car is occluded with another car in the scene and also occluded with some noise above the car due to BG adaptation transients of the GMM algorithm. As a result the blob appears much larger than the blob of a car and the reference method predicts the class label of the blob as a truck. In Fig.8.49(c), a van is occluded with a car at the left side of the image and also occluded with a human at the bottom side of the image. Therefore, the blob size of this occluded van is larger than its size in reality. Therefore, the class label prediction of the reference method is a truck for this blob. In the last example shown in Fig.8.49(d), a van has just entered the FOV of the cameras and has been occluded with a car. It can be argued that the respective blob could be designated as a car occluded with a van. However, in this thesis the ground truth labels are designated based on the class of the object, which is the closest to the camera(s). Hence, both methods misclassify the blob as if it were a car. Other than the last example, the proposed method classifies the first three examples correctly. For this reason, it is concluded that the radiometric features obtained with the source radiance estimate values, improves the overall classification accuracy with respect to the reference method.

8.6.5 Effects of Occlusion on Vehicle Classification

The moving vehicle detection clearly affects the classification performance. This is because the implemented method outputs occluded blobs from time to time. Therefore, in order to assess the performance of the classification method only, all the experiments have been repeated but this time the occluded blobs have not been used in the tests.

Firstly the effect of radiometric calibration and also the effect of atmospheric correction experiments are done without occluded samples. The first system that uses only *DN* features yields an accuracy of 98.94%. The system with the blackbody radiance features yields an accuracy of 99.06%, whereas the system with radiometric features has the highest accuracy, which is 99.2%. Therefore, in the case of an almost ideal vehicle detection method, the proposed radiometric features has a very high overall classification accuracy.

Secondly the reference system with the blob features and also the proposed system with both blob and radiometric features are investigated under the tests without occluded samples. For the single band measurement setup, the accuracy of the reference and proposed systems are 98.44% and 99.6%, respectively. The use of radiometric features is better than the use of blob features in terms of accuracy. Moreover, the use of radiometric features together with the blob features is even better again in terms of classification accuracy.

Thirdly, similar experiment is done for the dual band measurement setup. The accuracy of the reference and proposed systems are 98.16% and 99.34%, respectively. As expected, for the dual band measurement setup the proposed system with blob and radiometric features possesses higher accuracy than the reference system with blob features.

As a conclusion, all the results show that the use of the proposed radiometric features in vehicle classification significantly enhances the overall classification accuracy both in a practical system with occluded samples and also in a system with almost ideal vehicle detection block.

CHAPTER 9

CONCLUSION

A novel method that extracts radiometric features with atmospheric correction is proposed. The proposal is applied to vehicle classification both for a single and a dual band IR camera system. The proposed system has six major components that are registration, geometry estimation, atmospheric effects removal, data fusion, moving vehicle detection, and classification.

In this context, a literature search has been done for registration and a simple but effective registration process has been implemented. Since the effect of the atmosphere is different for each pixel in the frame, the geometric relation between the measured scene and the image plane is investigated with the proposed geometry estimation method so that the line of sight is estimated and used in the atmospheric correction. Then an atmospheric correction system is proposed, which considers the spectral characteristics of the detector, lens, and filter of the IR camera. In this process, an enhanced temperature calibration method is developed, too. The main contribution of the system is the accuracy due to the physical measurements and also the detailed consideration of the radiometric relationships. Literature search has been done for data fusion and a wavelet transform based data fusion process has been implemented. Similarly, a literature search has been done for vehicle detection process and a Gaussian mixture model based process is implemented. Finally, a literature search concentrated on SVM for the classification has been done. As the last but not the least contribution, the source radiance estimates are used in vehicle classification, which are the output of atmospheric correction process. To our knowledge this is the first attempt in the literature that radiometric quantities are used in vehicle

classification.

The proposed system is analyzed using two systems. First system consists of a single non-cooled LWIR camera and the second system consists of two cooled MWIR and LWIR cameras. With both systems, controlled measurements are taken. In the case of second system the measurements with both cameras are taken synchronously and the cameras are collocated. These records are analyzed in detail with the implemented methods. The temperature calibration method is shown to present a high accuracy, which is close to the NETD value of the cameras. Furthermore, it is shown that the atmospheric correction method removes the transmittance and path radiance effects from the measured data with the help of the proposed geometry estimation method and yields source radiance estimates. Then, the obtained radiance data are used in classification and it is seen that the proposed classification method increases the classification accuracies both for the first and second systems. As expected, the performance gain obtained with the dual band camera system is more than the one obtained with single band camera system at the cost of additional computational steps, which are registration and data fusion.

In the future, it is planned to improve the temperature calibration method by considering the reflected energy from the calibration device. Moreover, the output of geometry estimation can be utilized to extract features that represent physical quantities, e.g. physical width of a blob rather than pixel width. Another extension is to use different combinations of features in the classifier. For instance, the performance of a system with radiometric, blob, and also HOG features may be investigated. Finally, other classifiers than SVM, in the context of radiometric features, is an additional topic to be investigated.

REFERENCES

- [1] G.D. Sullivan, K.D. Baker, A.D. Worrall, C.I. Attwood, and P.M. Remagnino. Model-based vehicle detection and classification using orthographic approximations. *Image and Vision Computing*, 15(8):649–654, 1997. British Machine Vision Conference.
- [2] A.J. Lipton, H. Fujiyoshi, and R.S. Patil. Moving target classification and tracking from real-time video. In *Applications of Computer Vision, 1998. WACV '98. Proceedings., Fourth IEEE Workshop on*, pages 8–14, October 1998.
- [3] S. Gupte, O. Masoud, and N.P. Papanikolopoulos. Vision-based vehicle classification. In *Intelligent Transportation Systems, 2000. Proceedings. 2000 IEEE*, pages 46–51, 2000.
- [4] S. Gupte, O. Masoud, R.F.K. Martin, and N.P. Papanikolopoulos. Detection and classification of vehicles. *Intelligent Transportation Systems, IEEE Transactions on*, 3(1):37–47, March 2002.
- [5] H. Vceraraghavan, O. Masoud, and N. Papanikolopoulos. Vision-based monitoring of intersections. In *Intelligent Transportation Systems, 2002. Proceedings. The IEEE 5th International Conference on*, pages 7–12, 2002.
- [6] R.P. Avery, Y. Wang, and G. Scott Rutherford. Length-based vehicle classification using images from uncalibrated video cameras. In *Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on*, pages 737–742, October 2004.
- [7] Chung-Lin Huang and Wen-Chieh Liao. A vision-based vehicle identification system. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)*, volume 4, pages 364–367, Los Alamitos, CA, USA, 2004. IEEE Computer Society.
- [8] Stefano Messelodi, CarlaMaria Modena, and Michele Zanin. A computer vision system for the detection and classification of vehicles at urban road intersections. *Pattern Analysis and Applications*, 8(1-2):17–31, 2005.
- [9] Roya Rad and Mansour Jamzad. Real time classification and tracking of multiple vehicles in highways. *Pattern Recognition Letters*, 26(10):1597–1607, 2005.

- [10] Chengcui Zhang, Xin Chen, and Wei-Bang Chen. A PCA-based vehicle classification framework. In *Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on*, pages 17–17, 2006.
- [11] Xin Chen and Chengcui Zhang. Vehicle classification from traffic surveillance videos at a finer granularity. In Tat-Jen Cham, Jianfei Cai, Chitra Dorai, Deepu Rajan, Tat-Seng Chua, and Liang-Tien Chia, editors, *Advances in Multimedia Modeling*, volume 4351 of *Lecture Notes in Computer Science*, pages 772–781. Springer Berlin Heidelberg, 2006.
- [12] Jun-Wei Hsieh, Shih-Hao Yu, Yung-Sheng Chen, and Wen-Fong Hu. Automatic traffic surveillance system for vehicle tracking and classification. *Intelligent Transportation Systems, IEEE Transactions on*, 7(2):175–187, June 2006.
- [13] B. Morris and M. Trivedi. Robust classification and tracking of vehicles in traffic video streams. In *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*, pages 1078–1083, September 2006.
- [14] B. Morris and M. Trivedi. Improved vehicle classification in long traffic video by cooperating tracker and classifier modules. In *Video and Signal Based Surveillance, 2006. AVSS '06. IEEE International Conference on*, pages 9–9, November 2006.
- [15] T. Gandhi and M.M. Trivedi. Video based surround vehicle detection, classification and logging from moving platforms: Issues and approaches. In *Intelligent Vehicles Symposium, 2007 IEEE*, pages 1067–1071, June 2007.
- [16] D. Alonso, L. Salgado, and M. Nieto. Robust vehicle detection through multidimensional classification for on board video based systems. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 4, pages IV – 321–IV – 324, September 2007.
- [17] Guohui Zhang, Ryan P. Avery, and Yin Hai Wang. Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras. *Transportation Research Record: Journal of the Transportation Research Board*, 1993(1):138–147, January 2007.
- [18] N. Buch, J. Orwell, and S.A. Velastin. Detection and classification of vehicles for urban traffic scenes. In *Visual Information Engineering, 2008. VIE 2008. 5th International Conference on*, pages 182–187, July 2008.
- [19] Tuan Hue Thi, Kostia Robert, Sijun Lu, and Jian Zhang. Vehicle classification at nighttime using eigenspaces and support vector machine. In *Image and Signal Processing, 2008. CISP '08. Congress on*, volume 2, pages 422–426, May 2008.

- [20] Kostia Robert. Night-time traffic surveillance: A robust framework for multi-vehicle detection, classification and tracking. In *Advanced Video and Signal Based Surveillance, 2009. AVSS '09. Sixth IEEE International Conference on*, pages 1–6, September 2009.
- [21] J. Arróspide and L. Salgado. Region-dependent vehicle classification using PCA features. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 453–456, September 2012.
- [22] Lipchen A. Chan, Nasser M. Nasrabadi, and Vincent Mirelli. Multi-stage target recognition using modular vector quantizers and multilayer perceptrons. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 114–119, June 1996.
- [23] Lin-Cheng Wang, LipChen Chan, Nasser M. Nasrabadi, and Sandor Der. Combination of two learning algorithms for automatic target recognition. In *IEEE International Conference on Image Processing*, volume 1, pages 881–884, October 1997.
- [24] Lipchen A. Chan, Sandor Der, and Nasser M. Nasrabadi. Automatic target detection using dualband infrared imagery. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 6, pages 2286–2289, 2000.
- [25] Lipchen A. Chan, Sandor Z. Der, and Nasser M. Nasrabadi. Dual-band passive infrared imagery for automatic clutter rejection. In *IEEE International Conference on Image Processing*, volume 2, pages 463–466, September 2000.
- [26] Baoxin Li, Rama Chellappa, Qinfen Zheng, Sandor Der, Nasser M. Nasrabadi, LipChen Chan, and Lin-Cheng Wang. Experimental evaluation of FLIR ATR approaches - a comparative study. *Computer Vision and Image Understanding*, 84(1):5–24, 2001.
- [27] Bruce N. Nelson. Automatic vehicle detection in infrared imagery using a fuzzy inference-based classification system. *Fuzzy Systems, IEEE Transactions on*, 9(1):53–61, February 2001.
- [28] LipChen Alex Chan, Sandor Z. Der, and Nasser M. Nasrabadi. Improved target detector for FLIR imagery. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (2)*, pages 401–404, 2003.
- [29] LipChen Alex Chan, Sandor Z. Der, and Nasser M. Nasrabadi. Dualband FLIR fusion for automatic target recognition. *Information Fusion*, 4(1):35–45, 2003.
- [30] Syed A. Rizvi and Nasser M. Nasrabadi. Fusion techniques for automatic target recognition. In *IEEE Computer Society Applied Imagery Pattern Recognition Workshop*, pages 27–32, 2003.

- [31] Syed A. Rizvi and Nasser M. Nasrabadi. Fusion of FLIR automatic target recognition algorithms. *Information Fusion*, 4(4):247–258, 2003.
- [32] Xue Mei, Shaohua Kevin Zhou, and Hao Wu. Integrated detection, tracking and recognition for IR video-based vehicle classification. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 5, pages V–745–V–748, May 2006.
- [33] Mark W. Koch and Kevin T. Malone. A sequential vehicle classifier for infrared video using multinomial pattern matching. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on*, pages 127–133, June 2006.
- [34] S. Giompapa, A. Farina, F. Gini, A. Graziano, R. Croci, and R. Di Stefano. Naval target classification based on the confusion matrix. In *IEEE Aerospace Conference*, pages 1–9, March 2008.
- [35] Mark W. Koch and Kevin T. Malone. Vehicle classification in infrared video using the sequential probability ratio test. In Riad I. Hammoud, editor, *Augmented Vision Perception in Infrared*, Advances in Pattern Recognition, pages 63–84. Springer London, 2009.
- [36] Asif Mehmood and Nasser M. Nasrabadi. Anomaly detection for longwave FLIR imagery using kernel wavelet-RX. In *IEEE Computer Society International Conference on Pattern Recognition*, pages 1385–1388, 2010.
- [37] Çağlar Kavak and Seçkin Öz Saraç. Modeling atmospheric effects in the infrared scene generator based on OpenGL. In *IEEE 20th Signal Processing and Communications Applications*, pages 1–4, Fethiye, Muğla, Turkey, April 2012.
- [38] Çağlar Kavak and Seçkin Öz Saraç. Atmospheric effects modeling for the real time infrared scene generator. In *8th International IR Target and Background Modeling and Simulation Workshop*, Ettlingen, Germany, June 2012.
- [39] Çağlar Kavak and Seçkin Öz Saraç. Modeling atmospheric effects in the infrared scene generator with polynomial fitting. In *IEEE 21st Signal Processing and Communications Applications*, pages 1–4, Girne, Turkish Republic of Northern Cyprus, April 2013.
- [40] Çağlar Kavak and Seçkin Öz Saraç. Atmospheric effects modeling for the infrared scene generator with polynomial fitting. In *9th International IR Target and Background Modeling and Simulation Workshop*, Toulouse, France, June 2013.
- [41] T. M. Buzug, S. Schumann, L. Pfaffmann, U. Reinhold, and J. Ruhlmann. Skin-tumour classification with functional infrared imaging. In *Proceedings of the 8th IASTED International Conference Signal and Image Processing*, pages 313–322, Hawaii, USA, August 2006.

- [42] George Wolberg and Siavash Zokai. Image registration for perspective deformation recovery. In *Proc. SPIE*, volume 4050, pages 259–270, 2000.
- [43] Siavash Zokai and George Wolberg. Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations. *Image Processing, IEEE Transactions on*, 14(10):1422–1434, Oct 2005.
- [44] Helmuth Späth. Fitting affine and orthogonal transformations between two sets of points. *Mathematical Communications*, 9:27–34, 2004.
- [45] Hyunjin Yang, Joohwan Chun, Doo Chun Seo, and Jiyeon Yang. Radiometric calibration for MWIR cameras. In *Proceedings of SPIE Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXIII*, volume 8355, pages 1–6, 2012.
- [46] F. Meier, D. Scherer, J. Richters, and A. Christen. Atmospheric correction of thermal-infrared imagery of the 3-D urban environment acquired in oblique viewing geometry. *Atmospheric Measurement Techniques*, 4:909–922, 2011.
- [47] Alexander Berk, Gail P. Anderson, Prabhat K. Acharya, Lawrence S. Bernstein, Leon Muratov, Jamine Lee, Marsha Fox, Steve M. Adler-Golden, James H. Chetwynd, Jr., Michael L. Hoke, Ronald B. Lockwood, James A. Gardner, Thomas W. Cooley, Christoph C. Borel, Paul E. Lewis, and Eric P. Shettle. MODTRAN5: 2006 update. In *Proc. SPIE*, volume 6233, pages 62331F–62331F–8, 2006.
- [48] François Petitcolin and Eric Vermote. Land surface reflectance, emissivity and temperature from MODIS middle and thermal infrared data. *Remote Sensing of Environment*, 83:112–134, 2002.
- [49] Julia A. Barsi, John R. Schott, Frank D. Palluconi, and Simon J. Hook. Validation of a web-based atmospheric correction tool for single thermal band instruments. In *Proc. SPIE*, volume 5882, pages 58820E–58820E–7, 2005.
- [50] Bryan A. Franz, Sean W. Bailey, P. Jeremy Werdell, and Charles R. McClain. Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry. *Applied Optics*, 46(22):5068–5082, August 2007.
- [51] César Coll, Joan M. Galve, Juan M. Sánchez, and Vicente Caselles. Validation of Landsat-7/ETM+ thermal-band calibration and atmospheric correction with ground-based measurements. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(1):547–555, January 2010.
- [52] Kenneth Gade. A non-singular horizontal position representation. *Journal of Navigation*, 63:395–417, July 2010.
- [53] Berthold K. P. Horn and Robert W. Sjoberg. Calculating the reflectance map. A.I. Memo 498, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, October 1978.

- [54] Tomoo Mitsunaga and Shree K. Nayar. Radiometric self calibration. In *Proceedings of Computer Vision and Pattern Recognition*, volume 1, pages 374–380, Fort Collins, CO, 1999.
- [55] Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 Classes*, SIGGRAPH '08, pages 31:1–31:10, New York, NY, USA, 2008. ACM.
- [56] Ninh Duong and Michael Wegener. SensorVision radiometric equations version 2.2. DSTO Report DSTO-TN-0193, Defence Science and Technology Organisation, February 1999.
- [57] Claus Petersen and Niels Woetmann Nielsen. Diagnosis of visibility in DMI-HIRLAM. DMI scientific report no. 00-11, Danish Meteorological Institute, Copenhagen, Denmark, 2000.
- [58] İsmail Gültepe and G. A. Isaac. Visibility versus precipitation rate and relative humidity. In *12th Conference on Cloud Physics*, pages 55–62, Madison, WI, July 2006. American Meteorological Society.
- [59] A. Maynadier, M. Poncelet, K. Lavernhe-Taillard, and S. Roux. One-shot measurement of thermal and kinematic fields from Infra-Red Image Correlation (IRIC). In *ICEM 14-14th International Conference on Experimental Mechanics*, volume 6, pages 1–8, July 2010.
- [60] John C. Price. Calibration of satellite radiometers and the comparison of vegetation indices. *Remote Sensing of Environment*, 21:15–27, 1987.
- [61] Curtiss O. Davis, Jeffrey Bowles, Robert A. Leathers, Daniel Korwan, T. Valerie Downes, William A. Snyder, W. Joe Rhea, Wei Chen John Fisher, W. Paul Bissett, and Robert Alan Reisse. Ocean PHILLS hyperspectral imager: design, characterization, and calibration. *Optics Express*, 10(4):210–221, February 2002.
- [62] David D. R. Kohler, W. Paul Bissett, Robert G. Steward, and Curtiss O. Davis. New approach for the radiometric calibration of spectral imaging systems. *Optics Express*, 12(11):2463–2477, May 2004.
- [63] Joseph A. Shaw, Paul W. Nugent, Nathan J. Pust, Brentha Thurairajah, and Kohei Mizutani. Radiometric cloud imaging with an uncooled microbolometer thermal infrared camera. *Optics Express*, 13(15):5807–5817, July 2005.
- [64] Robert A. Leathers, Trijntje V. Downes, and Richard G. Priest. Scene-based nonuniformity corrections for optical and SWIR pushbroom sensors. *Optics Express*, 13(13):5136–5150, June 2005.

- [65] R. Guzmán, J. Meléndez, J. M. Aranda, Y. E. Essa, F. López, and J. L. Pérez-Castellanos. Measurement of temperature increment in compressive quasi-static and dynamic tests using infrared thermography. *Strain*, 45(2):179–189, 2009.
- [66] Gyanesh Chander, Brian L. Markham, and Dennis L. Helder. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113:893–903, 2009.
- [67] Nikolaus Anderson, Jeffrey Czaplak-Myers, Nathan Leisso, Stuart Biggar, Charles Burkhart, Rob Kingston, and Kurtis Thome. Design and calibration of field deployable ground-viewing radiometers. *Applied Optics*, 52(2):231–240, January 2013.
- [68] Pierre Tremblay, Louis Belhumeur, Martin Chamberland, André Villemaire, Patrick Dubois, Frédérick Marcotte, Charles Belzile, Vincent Farley, and Philippe Lagueux. Pixel-wise real-time advanced calibration method for thermal infrared cameras. In *Proc. of SPIE, Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXI*, volume 7662, pages 766212–766212–11, Orlando, Florida, April 2010.
- [69] Bo-Cai Gao, Marcos J. Montes, Ziauddin Ahmad, and Curtiss O. Davis. Atmospheric correction algorithm for hyperspectral remote sensing of ocean color from space. *Applied Optics*, 39(6):887–896, February 2000.
- [70] Kurt J. Thome. Absolute radiometric calibration of Landsat 7 ETM+ using the reflectance-based method. *Remote Sensing of Environment*, 78:27–38, 2001.
- [71] Luis Guanter, Karl Segl, Bernhard Sang, Luis Alonso, Hermann Kaufmann, and Jose Moreno. Scene-based spectral calibration assessment of high spectral resolution imaging spectrometers. *Optics Express*, 17(14):11594–11606, July 2009.
- [72] İbrahim Melih Çölova. An FPGA implementation of real-time electro-optic & IR image fusion. Master’s thesis, Middle East Technical University, Ankara, Turkey, September 2010.
- [73] Jian Yang, Jing yu Yang, David Zhang, and Jian feng Lu. Feature fusion: parallel strategy vs. serial strategy. *Pattern Recognition*, 36:1369–1381, 2003.
- [74] H. Li, S. Manjunath, and S. K. Mitra. Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 57(3):235–245, 1995.
- [75] Jamal Saeedi and Karim Faez. Infrared and visible image fusion using fuzzy logic and population-based optimization. *Applied Soft Computing*, 12:1041–1054, 2012.

- [76] Eduardo F. Nakamura, Carlos M. S. Figueiredo, and Antonio A. F. Loureiro. *Handbook of Algorithms for Wireless Networking and Mobile Computing*, chapter 37, pages 843–865. Chapman and Hall/CRC, 2005.
- [77] Serkan Özyürek. Image dynamic range enhancement. Master’s thesis, Middle East Technical University, Ankara, Turkey, September 2011.
- [78] F. Tsalakanidou, D. Tzovaras, and M .G. Strintzis. Use of depth and colour eigenfaces for face recognition. *Pattern Recognition Letters*, 24:1427–1435, 2003.
- [79] Seong G. Kong, Jingu Heo, Besma R. Abidi, Joonki Paik, and Mongi A. Abidi. Recent advances in visual and infrared face recognition - a review. *Computer Vision and Image Understanding*, 97:1103–135, 2005.
- [80] Jian Yang and Jing yu Yang. Generalized K-L transform based combined feature extraction. *Pattern Recognition*, 35:295–297, 2002.
- [81] Emrah Bala and Gözde Bozdağı Akar. Evaluation of image fusion and contrast enhancement algorithms for fusing LWIR and MWIR images. In *IEEE 16th Signal Processing and Communications Applications*, pages 1–4, Didim, Aydın, Turkey, April 2008.
- [82] Emrah Bala. An infrared fusion system. Master’s thesis, Middle East Technical University, Ankara, Turkey, August 2009.
- [83] Luca Bogoni and Michael Hansen. Pattern-selective color image fusion. *Pattern Recognition*, 34:1515–1526, 2001.
- [84] Bertan Günyel, Emrah Bala, and Gözde Bozdağı Akar. Fusion of long-wave and mid-wave infrared images. In *IEEE 15th Signal Processing and Communications Applications*, pages 1–4, Eskişehir, Turkey, June 2007.
- [85] Gonzalo Pajares and Jesús Manuel de la Cruz. A wavelet-based image fusion tutorial. *Pattern Recognition*, 37:1855–1872, 2004.
- [86] Zhang Feng, Liu Shang-qian, Wang Da-bao, and Guan Wei. Aircraft recognition in infrared image using wavelet moment invariants. *Image and Vision Computing*, 27:313–318, 2009.
- [87] Zexuan Zhu, Linlin Shen, Yiwen Sun, Shan He, and Zhen Ji. Memetic three-dimensional gabor feature extraction for hyperspectral imagery classification. In *Proceedings of the Third International Conference, ICSI 2012 on Advances in Swarm Intelligence Part I*, volume 7331, pages 479–488, Shenzhen, China, June 2012.
- [88] Demet Duman. Moving vehicle classification. Master’s thesis, Middle East Technical University, Ankara, Turkey, September 2013.

- [89] Chris Stauffer and W.E.L. Grimson. Learning patterns of activity using real-time tracking. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):747–757, August 2000.
- [90] Zoran Zivkovic. Improved adaptive gaussian mixture model for background subtraction. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)*. IEEE Computer Society, 2004.
- [91] Muhammet Pakyürek. A comparative evaluation of foreground / background segmentation algorithms. Master's thesis, Middle East Technical University, Ankara, Turkey, September 2012.
- [92] Zoran Zivkovic and Ferdinand van der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, 27(7):773–780, May 2006.
- [93] Ahmed Elgammal, David Harwood, and Larry Davis. Non-parametric model for background subtraction. In D. Vernon, editor, *European Conference on Computer Vision*, Lecture Notes in Computer Science, pages 751–767, 2000.
- [94] Christian Ruwwe and Udo Zölzer. GrayCut - object segmentation in IR-images. In *ICVS 2006 LNCS 4291*, pages 702–711, 2006.
- [95] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*, chapter 9. Prentice Hall Inc., Upper Saddle River, New Jersey 07458, second edition, 2002.
- [96] Robert M. Haralick and Linda G. Shapiro. *Computer and Robot Vision*, volume 1, pages 28–48. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, first edition, 1992.
- [97] J. Arróspide, L. Salgado, and J. Marinas. Hog-like gradient-based descriptor for visual vehicle detection. In *Intelligent Vehicles Symposium (IV), 2012 IEEE*, pages 223–228, June 2012.
- [98] S. Bileschi and L. Wolf. Image representations beyond histograms of gradients: The role of Gestalt descriptors. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1–8, June 2007.
- [99] Tuba Makbule Bayık. Automatic target recognition in infrared imagery. Master's thesis, Middle East Technical University, Ankara, Turkey, September 2004.
- [100] Demet Duman and Gözde Bozdağı Akar. Moving vehicle classification. In *IEEE 21st Signal Processing and Communications Applications*, pages 1–4, Girne, Turkish Republic of Northern Cyprus, April 2013.
- [101] Chris Harris and Mike Stephens. A combined corner and edge detector. In *In Proc. of Fourth Alvey Vision Conference*, pages 147–151, 1988.

- [102] Paul R. Beaudet. Rotationally invariant image operators. In *Proceedings of the International Joint Conference on Pattern Recognition*, volume 579, pages 579–583, 1978.
- [103] Tony Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):79–116, 1998.
- [104] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [105] John Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8(6):679–698, November 1986.
- [106] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*, chapter 10. Prentice Hall Inc., Upper Saddle River, New Jersey 07458, second edition, 2002.
- [107] Judith M. S. Prewitt. Object enhancement and extraction. In B.S. Lipkin, editor, *Picture Processing and Psychopictorics*, pages 75–149. Academic Press, New York, 1970.
- [108] J. Shi and C. Tomasi. Good features to track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, pages 593–600, June 1994.
- [109] Edward Rosten and Tom Drummond. Machine learning for high-speed corner detection. In Aleš Leonardis, Horst Bischof, and Axel Pinz, editors, *Computer Vision – ECCV 2006*, volume 3951 of *Lecture Notes in Computer Science*, pages 430–443. Springer Berlin Heidelberg, 2006.
- [110] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22(10):761–767, 2004. British Machine Vision Computing 2002.
- [111] Alfred Haar. On the theory of orthogonal function systems. *Mathematische Annalen*, 69(3):331–371, 1910.
- [112] Richard O. Duda and Peter E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15(1):11–15, January 1972.
- [113] D. H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.
- [114] M. Z. Mat Jafri and F. Deravi. Efficient algorithm for the detection of parabolic curves. *SPIE Vision Geometry III*, 2356:53–62, 1994.

- [115] Seçkin Öz Saraç and Gözde Bozdağı Akar. Breast segmentation in infrared images. In *IEEE 20th Signal Processing and Communications Applications*, pages 1–4, Fethiye, Muğla, Turkey, April 2012.
- [116] Gareth Loy and Alexander Zelinsky. A fast radial symmetry transform for detecting points of interest. In Anders Heyden, Gunnar Sparr, Mads Nielsen, and Peter Johansen, editors, *Computer Vision — ECCV 2002*, volume 2350 of *Lecture Notes in Computer Science*, pages 358–368. Springer Berlin Heidelberg, 2002.
- [117] Gareth Loy and Alexander Zelinsky. Fast radial symmetry for detecting points of interest. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(8):959–973, August 2003.
- [118] B.S. Manjunath, J.-R. Ohm, V.V. Vasudevan, and A. Yamada. Color and texture descriptors. *Circuits and Systems for Video Technology, IEEE Transactions on*, 11(6):703–715, June 2001.
- [119] D.G. Lowe. Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1150–1157 vol.2, 1999.
- [120] Herbert Bay, Tinne Tuytelaars, and Luc Gool. SURF: Speeded up robust features. In Aleš Leonardis, Horst Bischof, and Axel Pinz, editors, *Computer Vision – ECCV 2006*, volume 3951 of *Lecture Notes in Computer Science*, pages 404–417. Springer Berlin Heidelberg, 2006.
- [121] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893 vol. 1, June 2005.
- [122] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. ORB: An efficient alternative to SIFT or SURF. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2564–2571, November 2011.
- [123] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(10):1615–1630, October 2005.
- [124] O. Ludwig, Jr., D. Delgado, V. Goncalves, and U. Nunes. Trainable classifier-fusion schemes: An application to pedestrian detection. In *Intelligent Transportation Systems, 2009. ITSC '09. 12th International IEEE Conference on*, pages 1–6, October 2009.
- [125] Danian Zheng, Yannan Zhao, and Jiaxin Wang. Features extraction using a gabor filter family. In *Proceedings of the sixth LASTED International conference, Signal and Image processing*, Hawaii, USA, 2004.

- [126] F. Suard, A. Rakotomamonjy, A. Bensrhair, and A. Broggi. Pedestrian detection using infrared images and histograms of oriented gradients. In *Intelligent Vehicles Symposium, 2006 IEEE*, pages 206–212, 2006.
- [127] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, May 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [128] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*, pages 259–265. Wiley-Interscience, New York, USA, second edition, 2001.
- [129] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition*, pages 119–142. Academic Press, Elsevier Inc., MA, USA, fourth edition, 2008.
- [130] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition*, pages 198–203. Academic Press, Elsevier Inc., MA, USA, fourth edition, 2008.
- [131] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition*, page 573. Academic Press, Elsevier Inc., MA, USA, fourth edition, 2008.
- [132] Physics Dept., METU. Physics Dept. MeteoMonitor @ONLINE. <http://meteo.physics.metu.edu.tr/>, August 2014.
- [133] ITU. The radio refractive index: its formula and refractivity data. Recommendation ITU-R p.453-10, International Telecommunication Union, 2012.
- [134] ITU. Reference standard atmospheres. Recommendation ITU-R p.835-5, International Telecommunication Union, 2012.
- [135] Ali Mahmut Baldrige, Sacit Jane Hook, Cali Lu Grove, and G. Rivera. The ASTER spectral library version 2.0. *Remote Sensing of Environment*, 113(4):711–715, 2009.
- [136] E. Y. K. Ng. A review of thermography as promising non-invasive detection modality for breast tumor. *International Journal of Thermal Sciences*, 48:849–859, 2009.
- [137] William L. Wolfe. Radiation theory. In William L. Wolfe and George J. Zisis, editors, *The Infrared Handbook*, IRIA Series in Infrared & Electro-Optics, chapter 1, pages 1–41. The Infrared Information Analysis (IRIA) Center, Environmental Research Institute of Michigan, Ann Arbor, MI, fourth edition, June 1993.
- [138] R. Joro, A. L. Lääperi, P. Dastidar, S. Soimakallio, T. Kuukasjärvi, T. Toivonen, R. Saaristo, and R. Järvenpää. Imaging of breast cancer with mid- and

long-wave infrared camera. *Journal of Medical Engineering & Technology*, 32(3):189–197, May/June 2008.

- [139] N. Selvarasu, A. Nachiappan, and N. M. Nandhitha. Wavelet based abnormality extraction and quantification algorithm for thermographs depicting diseases in human. In *International Conference on Fiber Optics and Photonics*, Delhi, India, December 2008.
- [140] The MathWorks, Inc. MATLAB - The Language of Technical Computing @ONLINE. <http://www.mathworks.com/products/matlab/>, August 2014.

APPENDIX A

INFRARED THEORY

A.1 Planck's Radiation Law

Any object whose temperature is above absolute zero Kelvin emits radiation at a rate and with a distribution of wavelengths, λ . This wavelength distribution is dependent on the temperature of the object and its spectral emissivity, ϵ . The spectral emissivity, which may also be considered as the radiation efficiency at a given wavelength, is in turn characterized by the radiation emission efficiency based on whether the object is a blackbody, graybody or selective radiator. The blackbody is an ideal body. It is a perfect absorber that absorbs all incident radiation and as a consequence of Kirchoff's law is conversely a perfect radiator. According to Kirchoff's law; good absorbers are also good radiators and vice versa. This implies that a blackbody absorbs and emits the maximum theoretically possible energy at a given temperature. The spectral emissivity;

- $\epsilon(\lambda) = 1 \quad \forall \lambda$ for a blackbody
- $\epsilon(\lambda) = c \quad \forall \lambda$ where $0 < c < 1$ for a graybody
- $0 \leq \epsilon(\lambda) \leq 1$ for a selective radiator

The radiative power [$Wm^{-2}\mu m^{-1}$] as a function of wavelength and temperature is given by Planck's radiation law;

$$W(\lambda, T) = \frac{2\pi h_p c^2}{\lambda^5} \left[\exp\left(\frac{h_p c}{\lambda k T}\right) - 1 \right]^{-1} \quad (A.1)$$

where;

- h_p is Planck's constant $h = 6.6256 \times 10^{-34} Js$
- c is the velocity of light in vacuum $c = 2.9979 \times 10^8 m.s^{-1}$
- k is Boltzmann's constant $k = 1.38054 \times 10^{-23} W.s.K^{-1}$
- λ is the wavelength in μm
- T is the temperature in K

The absorptivity α of any material at a given temperature is equal to its emissivity ϵ at that temperature that is $\alpha(\lambda, T) = \epsilon(\lambda, T)$ [41, 136, 137].

The spectral blackbody radiance, L_{bb} of an object [$W.sr^{-1}.m^{-2}.\mu m^{-1}$] is computed with,

$$L_{bb}(\lambda, T) = \frac{W(\lambda, T)}{\pi} = \frac{2h_p c^2}{\lambda^5} \left[\exp\left(\frac{h_p c}{\lambda k T}\right) - 1 \right]^{-1} \quad (A.2)$$

when the object is assumed to radiate into a hemisphere in a homogeneous manner.

A.2 Wien's Displacement Law

The spectral density has a maximum at a certain wavelength for each temperature value. As the temperature increases, the wavelength of the maximum value decreases. This is the reason of the change in color of a glowing body as the temperature varies. Wien's displacement law states the inverse relationship between the wavelength of the peak of the emission of a blackbody and its temperature as [41]

$$T \lambda_{max} = 2898 \mu m K \quad (A.3)$$

A.3 Stefan-Boltzman Law

Stefan-Boltzman law helps to calculate the total radiant emittance. The infrared camera measures and images the emitted IR radiation from an object. The radiation is a function of the object surface temperature and so the temperature variations can

be calculated and displayed on the camera. This is done via the sensitive Stefan-Boltzman law,

$$\int_0^{\infty} \epsilon(\lambda, T)W(\lambda, T)d\lambda \propto T^4 \quad (\text{A.4})$$

The result of the integral is essentially the area below the Planck curve. As a consequence of Kirchoff's law, the radiation measured by the camera does not solely depend on the temperature of the object but is also a function of the emissivity [41].

APPENDIX B

INFRARED DETECTORS

The development of the infrared detectors goes back to the 1950s, where the aim was to use the detectors in military applications. IR detectors can be classified as thermal detectors and photon detectors based on the functional characteristics. Photon detectors rely on the interaction between the radiation quanta and the atoms of the material. On the other hand, thermal detectors rely on the temperature changes in the detector material. For temperature difference registering purposes, photon detectors are superior than the thermal detectors [138].

Photon detectors can be divided into five categories; namely photo conductive, photo voltage, photo emissive, photo magnetic, and quantum well. The most common materials used in the IR detectors are Gallium Arsenide (GaAs), Indium Antimonide (InSb), and HgCdTe. To minimize the thermal noise, which may exceed the detector signal caused by the photons, photon detectors are typically cooled to $77K$ to $80K$. In extrinsic semi-conductive detectors, even lower temperature values are required. Moreover, cooling process introduces a functional delay in the system. On the other hand, thermal detectors do not need cooling. There are two types of thermal detectors developed for industrial applications; bolometer and pyroelectric matrix detectors. These detectors approach the sensitivity of the photon detectors [138]. The most common detectors are given in the following sections.

B.1 Quantum Well Detectors

Quantum Well Infrared Photodetectors (QWIP) are one of the most sensitive detectors, which reach under $20mK$ NETD values. GaAs and similar materials are used in

QWIP detectors. Hence, QWIP detectors are thermally stable and radiation resistant. They have standard quality and high uniformity between detector elements. Their quantum efficiency¹ is between 1 – 12.5%. Moreover, these detectors can be adjusted to different wavelengths [138].

B.2 Photo Voltage Detectors

Photo Voltage (PV) detectors are the most advanced IR detectors, which reach under $18mK$ NETD values. InSb materials are used in PV detectors. PV detectors are maximally responsive to $3 - 5\mu m$ band and have quantum efficiency up to 90%. Thermal noise of PV detectors are decreased by cooling [138].

B.3 Micro Bolometer Detectors

Micro Bolometer (MB) detectors are non-cooled and so there is not any additional delay due to the cooling process. MB detectors have lower sensitivity than QWIP and PV detectors but almost an even response is achieved in a wide IR band. MB detectors may achieve $40mK$ NETD values when the f-value² is equal to one. HgCdTe and InSb materials are commonly used in MB detectors [138].

¹ Quantum efficiency is a compromise between detector material, pixel residual non-uniformity versus signal to noise ratio.

² f-value is the ratio between the focal distance and the aperture diameter.

APPENDIX C

GEOCENTRIC EARTH RADIUS

In [52], n -vector is given as,

$$\bar{n}^E = \begin{bmatrix} \sin(\lambda) \\ \sin(\mu) \cos(\lambda) \\ -\cos(\mu) \cos(\lambda) \end{bmatrix} \quad (\text{C.1})$$

where λ is the geodetic latitude, μ is the longitude and \bar{n}^E is the n -vector relative to the earth frame, E . In addition, the position vector of a point using the n -vector at an altitude h is given in [52] as,

$$\bar{p}_{EB}^E = \frac{b}{\sqrt{\sin^2(\lambda) + \frac{a^2}{b^2} \sin^2(\mu) \cos^2(\lambda) + \frac{a^2}{b^2} \cos^2(\mu) \cos^2(\lambda)}} \begin{bmatrix} \sin(\lambda) \\ \frac{a^2}{b^2} \sin(\mu) \cos(\lambda) \\ -\frac{a^2}{b^2} \cos(\mu) \cos(\lambda) \end{bmatrix} + h\bar{n}^E \quad (\text{C.2})$$

where a and b are the semi-major and semi-minor axes of the ellipsoid model corresponding to the equatorial and polar radius values of the earth, respectively. \bar{p}_{EB}^E is the position vector from the origin of frame E to the origin of body frame, B , relative to the earth frame. Then, if we set the altitude to zero and compute the $L2$ norm of the position vector, we get the following equation for the geocentric earth radius, R_{eff} .

$$R_{eff} = \|\bar{p}_{EB}^E(h=0)\|_2 \quad (\text{C.3})$$

$$R_{eff} = \sqrt{\frac{b^4 \sin^2(\lambda) + a^4 \cos^2(\lambda)}{b^2 \sin^2(\lambda) + a^2 \cos^2(\lambda)}} \quad (\text{C.4})$$

APPENDIX D

MULTIRESOLUTION THEORY

In multiresolution theory, images are represented and analyzed at more than one resolution. This is because, features that can not be detected at one resolution, may be detected easily at another resolution. If an object in an image is small in size or low in contrast, it is analyzed at higher resolutions. If an object is large in size or high in contrast, it is enough to analyze the object at a lower resolution. If both small and large objects, low and high contrast objects are present in an image, analyzing the image at several resolutions would be beneficial [139].

D.1 Wavelet Transform

Wavelet transform of an image produces four sub-band coefficients, which are approximation, horizontal detail, vertical detail, and diagonal detail coefficients. If $f(\cdot)$ is any square integrable function, the Continuous Wavelet Transform (CWT) of $f(t)$ with respect to a wavelet $\Psi(t)$ is defined as;

$$W(a, b) \triangleq \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \Psi^* \left(\frac{t-b}{a} \right) dt \quad (\text{D.1})$$

where a is a real number referred to as scale or dilation variable. b is a real number and represents the time shift or translation. $*$ denotes the complex conjugation [139].

CWT has a redundant representation of the signal since the entire support of $W(a, b)$ is not needed to recover the original $f(t)$. Hence, a non-redundant wavelet representation of the signal is;

$$f(t) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} d(k, l) \frac{1}{\sqrt{2}} \Psi(2^{-k}t - l) \quad (\text{D.2})$$

where $d(k, l)$ is called the Discrete Wavelet Transform (DWT). The equation does not involve a continuum of dilations and translations, instead it uses discrete values of a and b [139].

The Haar basis is obtained with a multiresolution of piecewise constant functions. The scaling function is $\phi = 1_{[0,1]}$. The impulse response of the filter $h[n]$ is,

$$h[n] = \begin{cases} \frac{1}{\sqrt{2}}, & \text{if } n = 0, 1 \\ 0, & \text{otherwise} \end{cases} \quad (\text{D.3})$$

The Haar wavelet is,

$$\Psi(t) = \begin{cases} -1, & \text{if } 0 \leq t < \frac{1}{2} \\ 1, & \text{if } \frac{1}{2} \leq t < 1 \\ 0, & \text{otherwise} \end{cases} \quad (\text{D.4})$$

where Haar wavelet is not well adapted to approximating smooth functions since it has only one vanishing moment [139].

APPENDIX E

INFRARED SIGNATURE ANALYSIS SOFTWARE

Infrared Signature Analysis software is an object-oriented software developed with Matlab[®] [140], which handles almost all the methods given in the thesis. IRSA is used to load the single or dual band binary image sequence files, apply registration if necessary, estimate the LOS geometry, compute all types of radiance values using radiometric calibration, apply data fusion if necessary, segment the images and finally extract all types of features mentioned throughout the thesis. At the last step, IRSA records all the extracted features into unique files together with the related meta data. The only method that is beyond the scope of IRSA is the classifier. As explained in the thesis multi-class SVM classifier is used and it is implemented as a separate software, which is capable of loading any feature recorded by IRSA and then yielding classification analysis results.

The Graphical User Interface (GUI) of IRSA is shown in Fig.E.1. In the Options menu the user has the option to choose the GUI language using Fig.E.2. When the user chooses the language as Turkish, GUI is seen as shown in Fig.E.3. In both Primary Frame Sequence and Secondary Frame Sequence fields of Fig.E.1, the respective frames are displayed. Clearly, in the single band case Secondary Frame Sequence field is not used. The sliders are used to display any desired frame in the sequence. Frame Time shows the local time of the computer on which the IR sequence is recorded. It shall be noted that as mentioned before signal generator is used to synchronize the cameras in the dual band measurement case, so the time difference between the respective frames is not an issue. Status Message field is used to inform the user if any unexpected behavior happens, e.g. the file containing GPS information cannot be found, etc. Load Camera Record File is used to load any binary sequence



Figure E.1: GUI of IRSA



Figure E.2: GUI of Language Selection

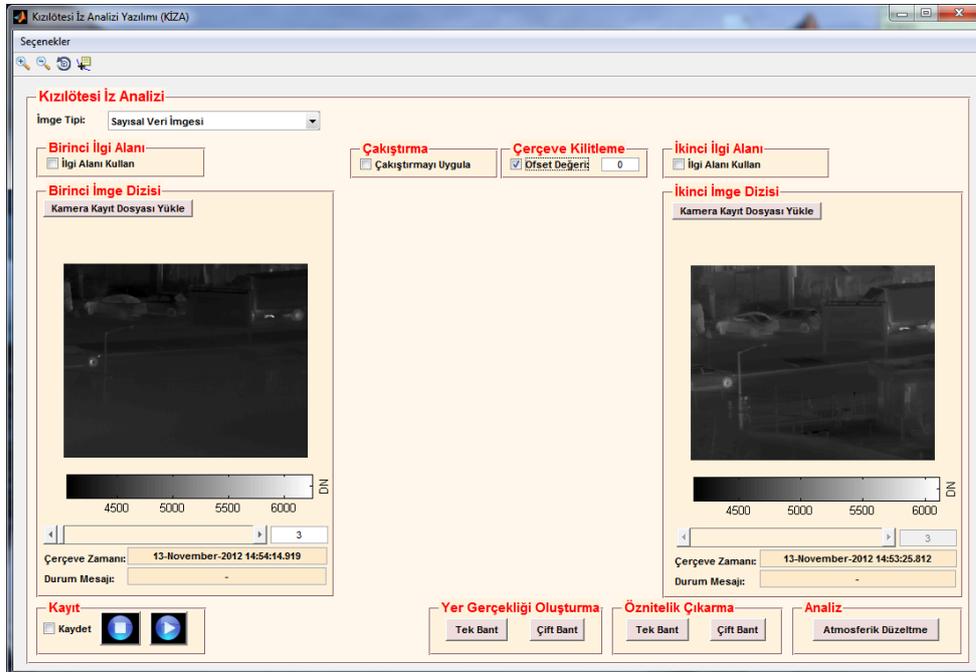


Figure E.3: GUI of IRSA in Turkish

file recorded with any IR camera used in the thesis. In the Image Type selection, the user chooses which type of the image to analyze. The image types are DN Image, Apparent Temperature Image, Apparent Radiance Image, Blackbody Radiance Image, Source Radiance Estimate Image, Range Values Image, Zenith Angle Values Image, Weighted Transmittance Map, Weighted Path Radiance Map, Fused Image, and Segmented Image. When the user chooses apparent temperature, GUI is shown in Fig.E.4. Similarly, when the user chooses segmented frame, GUI is shown in



Figure E.4: GUI of IRSA - Apparent Temperature Selection

Fig.E.5. In this case, Fused Frame Sequence field becomes visible and the segmented frame is displayed in this field. Moreover, the proposed normalization method for data fusion is selected in this field, too. In the Record field, the user chooses to play, pause, and stop the sequence. In addition, the user has the option to save the contents of the displayed image sequence. Registration field is used whether to apply registration or not in the case of dual band image sequences. Frame Lock field is used by the user again in the dual band case to offset any frame difference in between the sequences. Offset is required whenever more than a single number of frames during the start of the recording is dropped. Frame drops may be due to a reason, e.g. network congestion, etc. Atmospheric Correction button in the Analysis field runs the global analyses. When the button is pressed, all the respective image sequences are executed one by one and the global analysis results are recorded. Primary ROI and Secondary

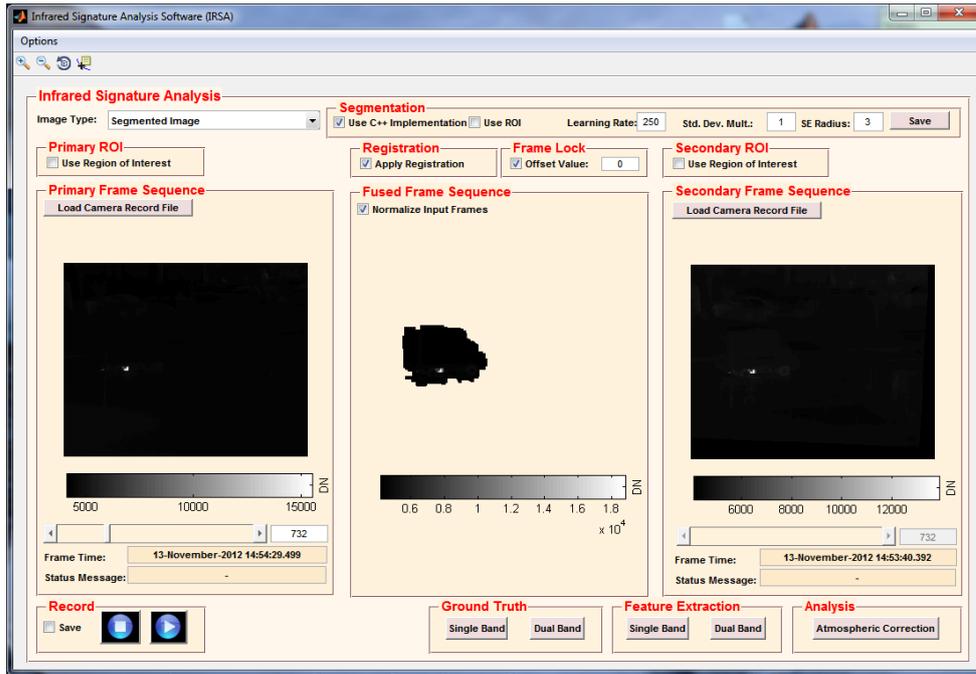


Figure E.5: GUI of IRSA - Segmented Image Selection

ROI fields are used in the global atmospheric correction analysis, to designate the ROI with asphalt region only in an interval of 100 frames. Feature Extraction field is used to record the extracted features. Single Band and Dual Band buttons execute the respective single and dual band image sequences one by one. Ground Truth field is used by the user to construct the ground truth labels manually for all the connected components corresponding to each extracted feature vector.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Öz Saraç, Seçkin

Nationality: Turkish (T.C.)

Date and Place of Birth: 25.01.1982, Gaziantep

Marital Status: Married with a child

Phone: +90 533 439 7766

Fax: +90 312 291 6004

EDUCATION

Degree	Institution	Year of Graduation
M.S	Bilkent University Electrical and Electronics Engineering	2008
B.S	Bilkent University Electrical and Electronics Engineering	2006

PROFESSIONAL EXPERIENCE

Year	Place	Enrollment
2008-2014	Tübitak Bilgem İltaren	Senior Researcher/Team Leader
2006-2008	Bilkent University	Research and Teaching Assistant

HONORS, SCHOLARSHIPS AND AWARDS

- Ranked 2nd in the quantitative area among 1.5 million students in the National University Entrance Exam of 2000
- Full scholarship for Overseas Undergraduate Study awarded by the Turkish Ministry of Education
- Full scholarship for Undergraduate Study awarded by Bilkent University
- Full scholarship for Master of Science Study awarded by Bilkent University
- Master of Science scholarship awarded by the Scientific and Technological Research Council of Turkey

PUBLICATIONS

International Journal Publications

- **Seçkin Öz Saraç** and Gözde Bozdağı Akar. Atmospheric effects removal for the infrared image sequences. *IEEE Transactions on Geoscience and Remote Sensing*, 2014. Revised draft submitted.
- **Seçkin Öz Saraç** and Gözde Bozdağı Akar. Effects of Radiometric Features on Vehicle Classification with Infrared Images. *IEEE Transactions on Intelligent Transportation Systems*. Under Preparation.
- **Seçkin Öz Saraç** and Ezhan Karaşan. Congestion window-based adaptive burst assembly for TCP traffic in OBS networks. *Photonic Network Communications*, 20(2):138–150, 2010.

International Conference Publications

- Çağlar Kavak and **Seçkin Öz Saraç**. Atmospheric effects modeling for the infrared scene generator with polynomial fitting. In *9th International IR Target and Background Modeling and Simulation Workshop*, Toulouse, France, June 2013.

- Çağlar Kavak and **Seçkin Öz Saraç**. Atmospheric effects modeling for the real time infrared scene generator. In *8th International IR Target and Background Modeling and Simulation Workshop*, Ettlingen, Germany, June 2012.

National Conference Publications

- Muhammed Yeşilkaya and **Seçkin Öz Saraç**. Cloud modeling in the infrared band. In *IEEE 21st Signal Processing and Communications Applications*, pages 1–4, Girne, Turkish Republic of Northern Cyprus, April 2013.
- Çağlar Kavak and **Seçkin Öz Saraç**. Modeling atmospheric effects in the infrared scene generator with polynomial fitting. In *IEEE 21st Signal Processing and Communications Applications*, pages 1–4, Girne, Turkish Republic of Northern Cyprus, April 2013.
- **Seçkin Öz Saraç** and Gözde Bozdağı Akar. Breast segmentation in infrared images. In *IEEE 20th Signal Processing and Communications Applications*, pages 1–4, Fethiye, Muğla, Turkey, April 2012.
- Çağlar Kavak and **Seçkin Öz Saraç**. Modeling atmospheric effects in the infrared scene generator based on OpenGL. In *IEEE 20th Signal Processing and Communications Applications*, pages 1–4, Fethiye, Muğla, Turkey, April 2012.
- **Seçkin Öz Saraç**. Rising sun reticle seeker with matched filter. In *IEEE 19th Signal Processing and Communications Applications*, pages 1185–1188, Ke-mer, Antalya, Turkey, April 2011.
- **Seçkin Öz Saraç** and Murat Akgül. Low pass filter cutoff frequency analysis and a new centroid method for the rosette scan seeker. In *IEEE 18th Signal Processing and Communications Applications*, pages 364–367, Diyarbakır, Turkey, April 2010.
- Mehmet Cihan Şahingil, **Seçkin Öz Saraç**, and Murat Akgül. Improved rising sun reticle. In *IEEE 18th Signal Processing and Communications Applications*, pages 700–703, Diyarbakır, Turkey, April 2010.