Signal and Image Processing for Enhanced Long Range Sensing

Improving long range sensing to counter developing threats

by

Fraser William Macfarlane

Centre for Signal and Image Processing

Department of Electronic and Electrical Engineering

University of Strathclyde, Glasgow

A thesis submitted in the fulfillment for the degree of

Doctor of Philosophy

February 6, 2026

Declaration of Authorship

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Fraser Macfarlane

February 6, 2026

Acknowledgements

Firstly, I would like to thank Professor Paul Murray for his support, mentorship, and friendship over the last 11 years since starting my university adventure, from small group tutor to PhD supervisor it has been quite a trip. And for helping me appreciate that "perfection" sometimes gets in the way of "very good". Secondly, I want to extend my gratitude to Professor Stephen Marshall who has always provided guidance and key insights over the course of the last 9 years since beginning my research career.

My work would not have been possible without the support of the ICASE award from the EPSRC and BAE Systems or, indeed, without data generously provided by both BAE Systems and DSTL via the University Defence Research Collaboration. Specifically I would like to thank Henry White for his technical expertise which helped to guide and focus my research. Additionally, I would like to thank Professor Gonzalo Arce and his group at the University of Delaware for providing a great deal of resources, guidance, and support in the latter stage of my PhD research - despite the collapse of international travel!

I would also like to thank everyone in CeSIP for their peerage and tutelage. To Connor Delaosa and David Vint for sharing this higher education journey with me. To Professor Stefan Weiss for his support and guidance over the years. To Dr Julius Tschannerl, Dr Andrew Young, Dr Andrew Campbell, Dr Jaime Zabalza, and Dr Stuart Bennett who all provided ample guidance, caffeination and ideas, often whilst putting up with my then youthful exuberance. To Dr He Sun, Dr Zhenyu Fang, and Dr Ha Viet Khanh for their cultural exchange, language lessons and spirited discussions. And also, my thanks to everyone in the HSI Lab for sharing ideas and always providing much needed discussion and troubleshooting and to everyone at the James Hutton Institute for encouraging me to get this across the line.

I would like to say thank you to all of my family and friends. While they might not understand exactly what it is that I do each day, that never stopped their encouragement. And finally, my loving gratitude to Agatha for her unwavering support, constant reassurance, and steadfast belief in me, even in times when I did not believe in myself.

Liònar beàrn mhòr le clachan beaga.

Great gaps may be filled with small stones.

ACC, MGM, & CMK

Abstract

Methods for rapidly identifying and discriminating between objects of interest and their background is key in many applications. With the increasing prevalence of hyperspectral cameras and the large datasets they produce, the need for efficient and interpretable tools for capturing, processing and analysing such data is imperative. The work presented in this thesis focuses on several techniques for improving upon existing methods of object and target detection in electro-optical imaging modalities.

This thesis is split into three distinct parts linked by the common theme of improving target detection in hyperspectral imagery. The first part investigates methods of extending the morphological Hit-or-Miss Transform for use in multivariate images as an efficient and explainable object detection algorithm. The second investigates methods of improving target detection in aerial hyperspectral imaging applications through the use of joint spatial and spectral dimensionality reduction. The final section investigates the use of simulated compressive sensing hardware for capturing a compressed representation of a scene and performing target detection and localisation without the need for any prior reconstruction.

The first contribution chapter focuses on extending a morphological approach to object detection, the Hit-or-Miss Transform. Traditionally constrained to single channel imagery, the proposed extension leverages the colour or spectral information afforded by multivariate data to detect objects based on their similarity to a reference in this high dimensionality space, as well as their size and shape. By incorporating the notion of percentage occupancy, the proposed method is made robust to noise and occlusion. Using various synthetic and natural images, the performance of the proposed Multi-Dimensional Percentage Occupancy Hit-or-Miss Transform is presented and compared with similar techniques showing that it can be more reliably applied to target detection tasks in multivariate data.

The second contribution chapter details a target detection pipeline using optimal spatial and spectral dimensionality reduction techniques. Target detection and classification is an important application of hyperspectral imaging in remote sensing as the increased spectral resolution allows for potentially greater distinction between target and non-target pixels in an imaged scene. Both the high redundancy, and sparsity of targets, in hyperspectral image data can be exploited. By eliminating pixels with known, non-target, characteristics prior to spectral dimensionality reduction, not only can the size of the data be decreased, but the representation of targets within the reduced domain is improved. The proposed approach, Joint

Spatio-Spectral Dimensionality Reduced Target Detection, achieves $> 95\%$ compression rates whilst preserving target detection performance in aerial hyperspectral images.

The third, and final, contribution chapter presents a method of performing target detection on data collected using Compressive Spectral Imaging without the need for prior reconstruction. This has the potential to save on the often costly and time intensive capture, storage, or transmission of highly redundant data, which is then largely discarded. Coded Aperture Snapshot Spectral Imaging captures compressed measurements of a scene via a known optical architecture based on the design of a coded aperture, as a result the behaviour of some desired target spectrum through the same architecture can be predicted. An approach for exploiting this, Coded Aperture Snapshot Spectral Imaging for Target Detection is presented as a baseline for performing target detection from compressed measurements and obtains promising results on synthetic and natural multispectral imagery.

The work presented in this thesis aims to improve upon target detection in hyperspectral imagery by exploiting the inherent sparsity in the data. Each contribution investigates one of many approaches to improving target detection, namely; efficient detection algorithms, optimal data compression, and imaging hardware. Demonstrating that this multifaceted problem, requires a multifaceted solution.

Contents

List of Figures

List of Tables

Acronyms/Abbreviations

ACE	Adaptive Coherence/Cosine Estimator	22
AD	Anomaly Detection	21
AUC	Area Under the Curve	72
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer	1
BG	Background	29
BUI	Built-up Index	3
CA	Coded Aperture	6
CASSI	Coded Aperture Snapshot Spectral Imaging	v, 6
CASSI-TD	Coded Aperture Snapshot Spectral Imaging for Target Detection	v, 6
CCASSI	Colour Coded Aperture Snapshot Spectral Imaging	201
CDF	Cumulative Distribution Function	100
CEM	Constrained Energy Minimisation	22
CHMT	Constrained Hit-or-Miss Transform	38
CNN	Convolutional Neural Network	27
CS	Compressed Sensing	6
CSI	Compressive Spectral Imaging	v, 1
DCT	Discrete Cosine Transform	26
DD-CASSI	Dual Dispersion CASSI	78
DD-CCASSI	Dual Dispersion Colour CASSI	205
DFT	Discrete Fourier Transform	85
DR	Dimensionality Reduction	3
EVD	Eigenvalue Decomposition	18
FAR	False Alarm Rate	72
FG	Foreground	29
FIR	Finite Impulse Response	22
FN	False Negative	72
FNR	False Negative Rate	72
FP	False Positive	42

FPCA	Folded Principal Component Analysis	18
FPR	False Positive Rate	97
GSD	Ground Sample Distance	170
HMT	Hit-or-Miss Transform	iv, 2
HSI	Hyperspectral Imaging	1
HYDICE	Hyperspectral Digital Imagery Collection Experiment	1
IC	Independent Component	19
ICA	Independent Component Analysis	19
JSSDR-TD	Joint Spatio-Spectral Dimensionality Reduced Target Detection	iv, 6
KL	Kullback-Leibler divergence	22
LRaSMD	Low-Rank and Sparse Matrix Decomposition	25
MCC	Matthews Correlation Coefficient	72
MDPOHMT	Multi-Dimensional Percentage Occupancy Hit-or-Miss Transform	iv, 5
MF	Matched Filter	22
MI	Mutual Information	20
MM	Mathematical Morphology	2
MNF	Maximum/Minimum Noise Fraction	18
MSI	Multispectral Imaging	1
NDSI	Normalised Difference Snow Index	21
NDVI	Normalised Difference Vegetation Index	3
$NDVI_{re}$	Normalised Difference Vegetation Index (red-edge)	20
NDWI	Normalised Difference Water Index	21
NIPALS	Non-linear Iterative Partial Least Squares	19
NIR	Near-Infrared	20
OSP	Orthogonal Subspace Projection	23
P_d	Probability of Detection	72
P_{fa}	Probability of False Alarm	72
PC	Principal Component	18
PCA	Principal Component Analysis	17
PD	Poisson Disc	80

PDF	Probability Density Function	100
PO	Percentage Occupancy	12
POHMT	Percentage Occupancy Hit-or-Miss Transform	5
PR	Precision-Recall	72
PUT	Pixel Under Test	21
R-CNN	Region Convolutional Neural Network	87
RENDVI	Red Edge Normalised Difference Vegetation Index	20
RGB	Red, Green, and Blue Colour Imagery	2
ROC	Receiver Operator Characteristic	71
ROI	Region of Interest	87
RXD	Reed-Xiaoli Detector	23
SAM	Spectral Angle Mapper	22
SD-CASSI	Single Dispersion CASSI	76
SD-CCASSI	Single Dispersion Colour CASSI	202
SE	Structuring Element	2
SHMT	Similarity Hit-or-Miss Transform	16
SID	Spectral Information Divergence	22
SNR	Signal-to-Noise Ratio	18
SPC	Single Pixel Camera	26
STBN	Spatiotemporal Blue Noise	84
SVD	Singular Value Decomposition	23
SVM	Support Vector Machine	19
TD	Target Detection	1
TN	True Negative	73
TNR	True Negative Rate	172
TP	True Positive	72
TPR	True Positive Rate	97
UHMT	Unconstrained Hit-or-Miss Transform	38
VAC	Void and Cluster	81
VD	Virtual Dimensionality	19
VI	Vegetation Index	3
VNIR	Visible and Near-InfraRed	52

WaLuDi	Ward's Linkage strategy using Divergence	65
WaLuMI	Ward's Linkage strategy using Mutual Information	65
WASP	Wildfire Airborne Sensor Program	1

Nomenclature

General Notation

x	Scalar
$\boldsymbol{x}$	Vector
$\boldsymbol{X}$	Matrix
X	Tensor
$\lvert x \rvert$	absolute value of scalar x
$\lVert \boldsymbol{x} \rVert$	norm of vector $\boldsymbol{x}$
$\lfloor x \rfloor$	floor value of scalar x
$\lceil x \rceil$	ceiling value of scalar x
μ	mean
σ	standard deviation
σ^2	variance
ρ	correlation

Hyperspectral Imaging

λ	wavelength
I	three-dimensional hypercube
$\boldsymbol{I}_{2D}$	two-dimensional representation of hypercube I
$\boldsymbol{I}_{\lambda}$	greyscale image at wavelength λ_i of hypercube I
$\boldsymbol{x}_i$	spectrum of i-th pixel in hypercube I
$\boldsymbol{I}_k$	greyscale line image representing k-th line of the hypercube I
$\boldsymbol{\mu}$	mean pixel vector
$\boldsymbol{\sigma}$	standard deviation pixel vector

Morphological Operators

$\boldsymbol{S}$	Structuring Element
$\check{\boldsymbol{S}}$	Reflected Structuring Element
$\boldsymbol{S}_{BG}$	Hit-or-Miss Transform Background SE

$\boldsymbol{S}_{FG}$	Hit-or-Miss Transform Foreground SE
$\varepsilon_{\boldsymbol{S}}(\boldsymbol{X})$	erosion of an image $\boldsymbol{X}$ by a structuring element $\boldsymbol{S}$
$\delta_{\boldsymbol{S}}(\boldsymbol{X})$	dilation of an image $\boldsymbol{X}$ by a structuring element $\boldsymbol{S}$
$\gamma_{\boldsymbol{S}}(\boldsymbol{X})$	opening of an image $\boldsymbol{X}$ by a structuring element $\boldsymbol{S}$
$\varphi_{\boldsymbol{S}}(\boldsymbol{X})$	closing of an image $\boldsymbol{X}$ by a structuring element $\boldsymbol{S}$
$\mathrm{HMT}_{\boldsymbol{S}}(\boldsymbol{X})$	Hit-or-Miss Transform of an image $\boldsymbol{X}$ by a composite structuring element $\boldsymbol{S} = [\boldsymbol{S}_{FG}, \boldsymbol{S}_{BG}]$

Probability Theory

$P(A)$	Probability of some event, A, occuring
$H(X)$	Shannon entropy of a variable X
$H(X,Y)$	join Shannon entropy between two variables X and Y
$I(X,Y)$	mutual information of two variables X and Y

Set Theory

$\mathbb{R}$	Set of real numbers
$\mathbb{Z}$	Set of integer numbers
$\mathbb{R}^N$	N-dimensional Euclidean space
$\mathbb{Z}^N$	N-dimensional discrete space
$A, B, ...$	sets $A, B, ...$
$a, b, ...$	elements of sets $A, B, ...$
$\varnothing$	Empty Set
$\forall$	for all
$\exists$	there exists
$\in$	element of
$\subseteq$	subset of
$\bigcup$	union
$\bigcap$	intersection
$\bigvee$	point-wise maximum (supremum)
$\bigwedge$	point-wise minimum (infimum)
$\mathrm{supp}(\cdot)$	the support, *i.e.,* the region defined by the non-zero elements, of $(\cdot)$
$\mathrm{card}(\cdot)$	the cardinal number, *i.e.,* number of elements, of $(\cdot)$

Chapter 1

Introduction

1.1 Motivation and aims

Object detection and classification from airborne platforms is a major challenge and active area of research within the discipline of signal and image processing. Hyperspectral Imaging (HSI) is a powerful tool for research in remote sensing among other areas and offers advantages over colour or Multispectral Imaging (MSI) sensors, predominantly in the wealth of information about an imaged scene it is able to collect. Hyperspectral cameras, such as the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) [1], or Hyperspectral Digital Imagery Collection Experiment (HYDICE) [2], capture information from hundreds of contiguous wavebands of the electromagnetic spectrum.

Other colour and MSI sensors including the Wildfire Airborne Sensor Program (WASP) [3], and open-access satellite imaging programmes such as the ESA's Copernicus [4], and the NASA/USGS Landsat [5], capture images at fewer, more precise, application-driven wavelengths. However, these sensors lack the spectral resolution, and resultant flexibility, offered by HSI. Hyperspectral sensors are also commonly used in prototype or lab-based systems for determining application specific wavelengths or in areas where small, narrowband, differences between spectra are required to be distinguished.

HSI allows for a vast quantity of information about a scene, and the objects contained within it, to be collected. As a result it has been used across a wide range of applications such as; industrial sorting [6], quality assessment in food, drinks and pharmaceuticals [7–9], smart agriculture [9–11], and defence and security [12, 13]. Such a range of use cases necessitates the development of tools to effectively and efficiently process and visualise this data.

With this in mind, three research themes were identified as critical areas of development in the field, when considering Target Detection (TD) in hyperspectral imagery, and were the focus of the work presented in this thesis. These themes were; updating classical, single channel, image processing techniques for use in high dimensionality data; exploiting spatial and spectral redundancies in HSI; and mitigating redundancy in image capture and performing detection in a compressed domain using Compressive Spectral Imaging (CSI). In this section how each of

these themes are linked to improving upon object and target detection is discussed.

1.1.1 Updating Morphological Image Processing Techniques for use in High-Dimensionality Data

Mathematical Morphology (MM) [14] is a simple yet powerful set of tools in image processing for analysing the size and shape of objects present in single channel images using concepts from set theory. The backbone of any morphological operation is Structuring Element (SE). By translating the SE(s) to each point in an image and applying set operations such as unions and intersections or some combination, objects can be grown, shrunk, or removed. Achieving the desired outcome is not only dependent on the morphological operation used, but also on the optimal design of the SE(s) used.

One technique in particular, the Hit-or-Miss Transform (HMT) [15], is used for probing an image with two such SEs in order to detect objects of known size and shape. As with all morphological operations, a major drawback of the HMT is the lack of a generally accepted extension to images with more than one channel. This extension to multi-channel images, such as traditional Red, Green, and Blue Colour Imagery (RGB), MSI, and HSI, as well as volumetric images, such as computed tomography and magnetic resonance imaging, is desirable. By treating each channel individually, as is the case for some proposed extensions of morphology, the relationship between channels is lost [16, 17]. A multi-channel approach would be able to exploit these correlations, or indeed the inherent multi-dimensionality, that exists in such images, *i.e.,* the colour or spectrum of objects in an RGB or hyperspectral image or the volume and contours of objects within volumetric images. A general multi-channel HMT would allow the benefits of fast processing, simplicity, and interpretability to object detection in such images, without the requirement for extensive annotated training data; as would be the case when using a machine/deep learning approach. In defence and security applications, the explainability and interpretability of these operations is also desirable due to the lack of a so-called "black box". The development of one such proposed technique for a multi-channel HMT as an efficient method of performing object detection is described in this thesis.

1.1.2 Exploiting Redundancies in Hyperspectral Imaging

In addition to spatial object detection through morphological and deep-learning methods, TD methods have been designed that can take advantage of the high-resolution spectral information provided within hyperspectral images. Although HSI offers a large number of observations about

an underlying scene, this high dimensionality can degrade the performance of signal and image processing techniques. In this thesis, dimensionality refers to the number of elements in a given feature vector.

Hughes Phenomenon, or *"the curse of dimensionality"*, can become a factor; *i.e.,* as the number of dimensions in a dataset increases, the predictive power of classifiers initially increases accordingly; however, beyond some threshold their performance begins to degrade [18, 19]. This can be attributed to a number of factors. As the number of features, or dimensions, increases, this can lead to increased sparsity where observations become more distributed in a high dimensional space and the distances or inter-relationships between them become increasingly similar and therefore less informative.

Hughes phenomenon can be avoided through feature extraction or feature selection techniques to reduce the inherent dimensionality of data by inferring new features or retaining only the most important. In particular, feature transformations and Dimensionality Reduction (DR) techniques have been shown to mitigate the effects of Hughes phenomenon [20, 21]. DR seeks to represent high dimensionality data, in a lower dimensionality space. In the case of HSI, given a set of *a-priori* parameters, optimal for the scenario, an optimal embedding can be obtained. However, this is more likely to be carried out in an *a-posteriori* manner using the image data to obtain this optimal embedding. In addition to simplifying imagery using spectral measures, Vegetation Indices (VIs) [22], and other remote sensing indices, offer a technique for reducing the spatial dimensionality of data. Using techniques such as the Normalised Difference Vegetation Index (NDVI) or the Built-up Index (BUI) not only provides additional information about an imaged scene, but can determine whether or not the underlying pixel is relevant to the application. In this thesis, combining both spatial and spectral DR for the purpose of improving the efficiency and efficacy of TD applications are investigated.

1.1.3 Mitigating Redundancy in Image Capture using Compressive Spectral Imaging

Capturing the entire scene, only to then compress and selectively discard unnecessary information is wasteful in terms of time, energy, storage, and compute resources. As a result, methods of limiting these overheads are of interest. CSI systems offer a method for capturing compressed representations of a scene which can then reconstructed at a later date. Such systems allow for DR and optimal compression to be achieved at the hardware level given proper design. As sufficient information is retained in the sparse representation of an imaged scene using CSI such that an approximation of the original image can be made, investigation

into whether this information is sufficient for spectral TD is undertaken. Such an approach avoids the need for reconstruction prior to further processing, and would further simplify the processing pipeline for such approaches. The final chapter of this thesis involves developing CSI-based methods to perform TD based on compressed measurements.

1.1.4 Research Aims

The work presented in this thesis will focus predominantly on developing efficient algorithms for performing TD in spectral imagery as well as DR and compression. The following objectives were identified and investigated as part of this thesis:

1. Develop novel morphological operations which can be applied to multi-channel images in order to rapidly and efficiently detect objects based on known size, shape, and colour parameters without the need for large training samples.
2. Investigate efficient compression techniques for use in hyperspectral TD applications in order to reduce the size and complexity of an image and determine how such compression affects performance.
3. Investigate the use of compressed sensing techniques and determine their efficacy in compressive TD tasks - capturing compressed representations of a scene and identifying spectral features of interest without the need for reconstruction.

The contributions and techniques developed after pursuing these areas of investigation are introduced in more detail in the next section.

1.2 Novel Contributions

The work presented in this thesis investigates target/object detection, efficient dimensionality reduction, and compression, as highlighted in the previous section, by developing algorithms and application-driven novel image analysis approaches. The novel contributions of this work are:

1. **A new definition of the morphological Hit-or-Miss Transform for use in colour and other multichannel images**

   Mathematical Morphology presents a powerful set of tools for analysing imagery based on size and shape information, or morphology. However, morphological techniques are mostly limited to single-channel greyscale images and treat an image as a topological surface. One morphological approach, the Hit-or-Miss Transform, identifies objects of interest based on local regions which fit constraints. To take advantage of the information provided by colour and spectral data in addition to the size and shape of objects, a new definition for a colour HMT, the Multi-Dimensional Percentage Occupancy Hit-or-Miss Transform (MDPOHMT), is presented.

2. **An extension of the Percentage Occupancy Hit-or-Miss Transform for use in multivariate imagery**

   The original definitions of MM are rigid in that morphological operators are based on *minimum* and *maximum* filters, with no accommodation for the presence of noise or other degrading effects commonplace in images. This limits the use of such operators in detecting noisy or non-ideal objects such as when targets are incomplete or occluded. These harsh limits can be relaxed through the use of rank order filters as has already been applied to binary and greyscale HMTs. In this thesis, rank order-based relaxations have been incorporated and expanded to improve the noise robustness of contribution 1 for detecting objects in multivariate imagery.

3. **An innovative technique for estimating the parameters of the novel MDPOHMT**

   Another improvement to the original definitions of morphology, specifically the HMT and the Percentage Occupancy Hit-or-Miss Transform (POHMT), is the ability to define the relaxation parameters of the POHMT for a set of example objects. Like the extension of the HMT to the POHMT, this parameter setting technique can be modified for use in multivariate images and a set of design tools are displayed in this thesis.

4. **A method for performing spatial dimensionality reduction in hyperspectral images using remote sensing indices for application to Target Detection.**
   Hyperspectral images are highly redundant in the spectral dimension due to the high correlation between subsequent sets of bands. However, in certain applications, such as TD, there exists large amounts of spatial redundancy where spectra of interest do not occur and this absence can be measured. Using remote sensing indices, such as a VI, non-target pixels can be removed from images in order to simplify and improve subsequent spectral dimensionality reduction and TD. This spatial masking was used to improve the efficiency of the novel multivariate HMT, contribution 1, by greatly reducing the set of locations required to probe an image, and in contribution 5 to inform which subset of pixels should be considered for DR and TD.

5. **A pipeline for combining spatial and spectral dimensionality reduction for Target/Anomaly Detection and Classification techniques**
   The techniques developed and findings from contribution 4 were then formalised into the Joint Spatio-Spectral Dimensionality Reduced Target Detection (JSSDR-TD) pipeline which combines optimal spatial and spectral DR to improve the efficiency of performing TD on hyperspectral images. Previously, the spatial DR component acted purely as a binary mask on already spectrally dimensionality reduced imagery. Instead, this spatial masking can be used to refine the domain of an image by removing non-target pixels prior to any compression into a dimensionality reduced subspace, improving the detection due to the suppression of any non-target signals and simplifying the DR process by removing the need to account for them in a optimal subspace.

6. **A novel method for detecting targets of interest in compressively sensed images**
   Compressed Sensing (CS) has been used to great effect in capturing signals and images using sub-Nyquist sampling. CS techniques require often complex and resource intensive mathematical optimisation to create a reconstructed signal from these compressed measurements. Using simulated Coded Aperture Snapshot Spectral Imaging (CASSI) architectures, a TD method for identifying spectra of interest in compressed representations of an image scene are presented in this thesis. Additionally, various Coded Aperture (CA) generation methods are investigated for use in this proposed Coded Aperture Snapshot Spectral Imaging for Target Detection (CASSI-TD) approach.

7. **A set of discretised MATLAB models for simulating the behaviour of CASSI architectures**

   Whilst developing methods for performing TD using compressed CASSI measurements, it was necessary to construct a series of simple models for simulating various CASSI architectures. MATLAB models simulating both single and dual dispersion CASSI for both binary and colour architectures have been developed as part of this work. Additionally, MATLAB algorithms based on existing techniques for CA generation have been implemented to facilitate these simulations.

1.2.1 Contribution Publication Summary

- The findings from contributions 1-3 were initially presented at EUSIPCO 2018 in Rome and HSI 2018 (now SPIE Photonex) in Coventry. Following a number of additional experiments, the results and technique were published in Mathematical Morphology: Theory and Advances as detailed in Appendices C.1 and C.2.
- Contribution 4 formed part of the work detailed in conference papers presented at HSI 2018 in Coventry and SPIE Defence and Commercial Sensing 2019 in Baltimore (Appendix C.2).
- The findings from contributions 4 and 5 were initially presented at SPIE DCS 2019 in Baltimore and the extended experiments and results were published in a special issue of MDPI Remote sensing: Feature Extraction and Data Classification in Hyperspectral Imaging as detailed in Appendices C.1 and C.2.

1.3 Thesis Outline

The remainder of this thesis is set out in the following chapters.

- Chapters 2 and 3 contain a review of the relevant literature and theoretical background to each of the concepts presented in the rest of this thesis respectively.
- Chapter 4 details the development of the Multi-Dimensional Percentage Occupancy Hit-or-Miss Transform and displays the results of its use in various object detection scenarios. The design tool for selecting the optimal operational parameters and the effects of relaxation for noise robustness are also presented.
- Chapter 5 presents the novel Joint Spatio-Spectral Dimensionality Reduced Target Detection pipeline for performing efficient TD in hyperspectral imagery. A comparison of DR and TD algorithms for hyperspectral data analysis is also presented.
- Chapter 6 outlines the simulation of both binary and colour Coded Aperture Snapshot Spectral Imaging hardware and displays the results of performing TD on compressed representations of multispectral images. The design of the simulated CASSI hardware is presented and multiple CA generation techniques which underpin the TD are displayed and compared.
- In Chapter 7, the work presented in this thesis is summarised with a examination of the novel contributions and their limitations. Additionally, suggestions for improvements and future work not considered in this thesis are discussed.
- Appendices A and B contain the datasets used in the works presented and supplementary results from each of the contributions in this thesis respectively.
- Appendix C contains the author's published papers during the PhD.

Chapter 2

Literature Review

2.1 Introduction

Hyperspectral Imaging, is commonly used in remote sensing applications as well as various other research areas. HSI offers a vast quantity of spectral information about the scene it captures and with this comes the need for efficient and effective processing. The main focus of this thesis is to investigate methods of efficient HSI-based object detection and classification within remote sensing applications and the development of novel techniques to facilitate this. In this chapter, the existing literature concerning various image processing techniques related to the novel contributions presented throughout this thesis are reviewed. Beginning with MM (Section 2.2) which is utilised and extended to act as an efficient object detection method in multichannel images. This is followed by DR and TD (Sections 2.3 and 2.4) both of which are routinely used in the analysis of hyperspectral imagery. Finally, an overview of CS theory is presented to discuss how it can be used to efficiently capture compressed representations of the highly redundant information present in hyperspectral imagery prior to any further transmission, processing, or reconstruction.

2.2 Mathematical Morphology

In this section, the origins, extensions, and recent advancements of Mathematical Morphology are reviewed and discussed in relation to the morphological contributions covered later in this thesis. The original binary, and later general single channel, definitions are covered first as well as more recent work on incorporating forms of noise robustness. The various efforts to extend the domain of Mathematical Morphology to colour and multivariate images are also reviewed here.

2.2.1 Binary Morphology and Greyscale Extensions

Mathematical Morphology was first developed in the 1960s by Georges Matheron [23] and Jean Serra [14, 15, 24] and was later extended by Heijmans [25]. Initially developed for the analysis

of geological data and the structure of materials, specifically the quantitative analysis of the quality of iron ore [26], MM has evolved into a powerful set of non-linear image processing techniques used for analysing the shape and structure of objects within an image, *i.e.,* their morphology. It is defined as a mathematical technique due to its origins in set theory, integral geometry, and lattice theory.

The original definitions of morphological operations [14, 23] dealt solely with binary images, or sets, as elements in the images of porous media they were formalised to analyse could either belong to a pore or the matrix surrounding them [27]. This idea is analogous to two complementary sets in mathematics and using simple set theory operations such as unions, intersections, complementation, and translation allows for fast analysis of objects of potential interest. As such, a set of image-based operations was defined. MM seeks to apply the mathematical concepts of set theory, specifically lattice theory, to various image analysis tasks such as object or edge detection [28–30], segmentation [31–33], and image filtering and de-noising [27, 34]. Whilst originally intended for use in binary images, morphological techniques have since been extended for application to greyscale images [35–37] and more recently an active focus has been on further extending mathematical morphology for use in multivariate images [38–40].

The basis of any morphological operation is the SE. The SE is simply a small set, or filtering window, which is used to probe an image under analysis and define a region, or support, for some particular morphological operation to take place [27]. SEs can traditionally be described as being either *flat* or *non-flat* with flat SEs being simple binary masks which define whether a pixel is included in the operation or not. Non-flat SEs however, do not only define the bounded region of a morphological process, they also take on some greyscale value which is added, or subtracted, to the associated pixel. This addition, or subtraction, can be used to alter a result as they introduce new values that would not be present using a flat SE.

The two elementary morphological operators are erosions and dilations [27]. The erosion is performed by probing a query image with an SE at each pixel and can be defined as the infimum, or minimum value, within the support of the SE. Similarly, the dilation is defined as the supremum, or maximum value, within the support of a probing SE. Erosions and dilations can in turn be used to define additional operations such as opening and closing which are performed by applying an erosion followed by a dilation and vice versa.

The morphological HMT [14], in its initial definition for binary images, is implemented by combining two erosions in the supports of two, non-intersecting, SEs. These SEs are used to probe the foreground and background of an image respectively and returns areas where

the foreground SE matches the image and the background SE simultaneously fits around any potential features of interest. The mathematical derivation of each of these operators as well examples of their uses are presented in Section 3.2.

Whereas binary images can be viewed as two complementary sets with elements 1 and 0 representing the foreground and background of that image respectively, greyscale images can take on a range of values dependant on their bit-depth. The elementary operators of erosion and dilation, and the openings and closings derived from them, are easily implemented in greyscale images due to them being increasing transforms [27].

Unlike the other common MM operations, the HMT is not an increasing transform and as a result any further extension to greyscale images is not straightforward [41], with few authors considering any possible extension [42]. The predominant difficulty in extending the HMT for use in greyscale images is the necessity to take the complement of an image in order to define the HMT (as shown and discussed later in Section 3.2). This resulted in multiple definitions of the HMT for greyscale images in [36, 43–45] with a review of each and a resultant unified theory being developed in [42]. Instead of eroding both the image itself by the foreground SE and eroding the complement of the image by the background SE, the dual operation of dilation was incorporated and the transform could be defined in terms of an erosion of the foreground and a dilation of the background [27, 46].

One main drawback of MM in general, and the HMT specifically, is the susceptibility of such methods to fail in the presence of noise. A single noisy pixel in either the background or foreground of the image, within the support of a probing SE, can drastically change the output of a morphological operation. For example, in the case of the HMT, a single pixel in the support of the background SE which has a value greater than, or equal to, those in the support of the foreground SE can prevent the detection of features that would otherwise be detectable. An extensive review of methods which provide some element of noise robustness in HMTs is given in [46] with three main categories being identified. The first is the use of a de-noising or other pre-processing steps in order to aid the fitting of any SEs. The second involves the alteration of any given SE to account for noise. The third and final approach is to alter the definition of the transform in order to relax the strict fit required by the HMT.

Removing sufficient noise from an image without also altering the fundamental features of the object to be detected is often a challenging endeavour. Similarly, creating individual SEs in a bespoke manner to account for the vast range of noise characteristics is an infeasible task [47]. For these reasons there have been many attempts in the literature, as outlined in [46], where authors take the third approach, proposing alternative methods of implementing morphological

operators by relaxing the required fit of a given SE.

Rank order filters and the fundamental operations of MM are intrinsically linked [36]. The result of a rank order filter is defined as the k^{th} element of an ordered set of values. The median filter, the output of which is calculated as the value corresponding to middle rank of an ordered set within the support, is arguably the most common example of this family of filters. By using rank order filters, alternative definitions for morphological operators can be developed. The result of an erosion is defined as being equivalent to the minimum value within the support of some SE, $\boldsymbol{X}$, or alternatively the value at rank $k = 1$ when the elements within such a support are ordered. Likewise, a dilation is the maximum value within the support of an SE and can therefore be implemented using a rank $k = \text{card}(\boldsymbol{X})$ filter, where $\text{card}(\boldsymbol{X})$ defines the number of elements, or cardinality, in the support of $\boldsymbol{X}$. As described theoretically in [47] and demonstrated in [48], rank order filters can be used to relax the hard constraints of minimum ($k = 1$) and maximum ($k = \text{card}(\boldsymbol{X})$) filters used in erosions and dilations respectively.

A measure for defining how well some SE fits when considering the HMT of an image in the presence of noise is defined in [47]. This subsequently gave rise to the notion of Percentage Occupancy (PO) [49] and the POHMT which presented a generalisation of multiple greyscale HMTs with the addition of explainable noise robustness. The calculated PO values can be used to inform a pair of rank order filters which in turn can be used to detect features of interest despite noise in an image. Additionally, a set of design tools was developed to inform this process [50].

2.2.2 Colour and Multivariate Extensions of Mathematical Morphology

The notion of order is essential for understanding and implementing morphological operations as it allows for the description of a set of values as an ordered list. From this, morphological operations can be derived from the ordered set using the extrema, allowing for the definition of both the infimum and supremum necessary for performing erosion and dilation respectively. An ordering relation is a way of sorting elements in a set by comparing pairs of elements and determining their relationship to one another, *e.g.*, for numbers: $\leq$ and $\geq$, for sets: $\subseteq$ and $\supseteq$, or even for concepts such as shapes fitting within or containing each other. In the case of morphology these elements are the pixel values within the support of some SE. The type of order obtained is dependent on the adherence to the following criteria for all elements x, y and z in a given set, X, for some relation, R, in this case '$\leq$':

1. Reflexivity: $x \; R \; x$,
2. Anti-symmetry: $x \; R \; y$ and $y \; R \; x$, only if $x \; = \; y$,
3. Transitivity: if $x \; R \; y$ and $y \; R \; z$, then $x \; R \; z$,
4. Connexity or triochotomy: exactly one of $x \; < \; y$, $x \; = \; y$, or $x \; > \; y$, $\forall \; x, y \; \in \; X$.

A pre-ordered set is one which satisfies both the reflexivity and transitivity properties, whilst a partial order additionally satisfies the anti-symmetry property. In a partial ordering, not all elements $x \in X$ are required to have any relation with every other $y \in X$. A totally ordered set meanwhile, requires a stricter form of the anti-symmetry principle called *trichotomy* [27], where every pair of elements, x and y, in the set must satisfy exactly one of the given relations. Defining a total order in binary or greyscale images is fairly trivial as there are obvious extrema and a natural order to each of the values present.

Unlike in single channel images, there is no obvious or unambiguous way of ordering multivariate data. There have been, however, multiple studies investigating methods of, as well as the necessity for, introducing ordering relations in multivariate data. According to Barnett [51], there are four general sub-categories that multivariate ordering relations can belong to; marginal, partial, conditional, and reduced orderings.

The first, marginal, or *m-ordering*, is concerned with treating each vector as a set of scalars or, in the case of image analysis techniques, processing each channel in an image separately. By ordering each channel separately, the overall context is lost and this can have the unwanted effect of introducing colour vectors which were not present in the original image, also known as the false colour problem [17, 52].

The second method is partial, or *p-ordering*, which separates the vectors into groups which can then be distinguished by some measure of rank or extremeness. For example, this can be carried out using convex hulls [38] which encompass all of the samples in multidimensional space. The samples on the perimeter of the convex hull can then be discarded and the process repeated. The discarded samples give a measure of extremeness at each iteration and thus an order can be derived. Importantly, a partial ordering is distinct from a partial order defined by the characteristics discussed previously.

The third method is conditional, or *c-ordering*, which involves ordering one of the marginal sets based on some condition. This can be repeated on subsequent sets if the previous result is inconclusive *e.g.,* first ordering the red channel in an RGB image followed by the green then blue if the red intensity values are repeated and a true order cannot be determined. *c-ordering*, or lexicographical ordering as it is sometimes referred to in literature, has the disadvantage that

the infimum and supremum defined are wholly dependant on the order in which the channels are processed [53]. This may introduce biases and should only be used where there is a known priority among image channels [52] that is invariant to, or augmented by, these biases. For example, a lexicographical approach in the HSV colour space is presented in [29] where the intensity values are considered most important over saturation and hue with colours ordered following this convention.

The final method, and the one employed in the proposed extension to the HMT in Chapter 4, is reduced, or *r-ordering*, which involves measuring the vector against some reference vector or by projecting it into a reduced space. This creates a scalar value based on the disparity of each pixel vector to the reference, which can more easily be sorted.

Both *r-ordering* and *p-ordering* rely on a pre-ordering of data meaning they lack the anti-symmetry characteristic [54]. In the case of *r-ordering*, this is caused by the potential for two distinct vectors, $a \neq b$, to be equidistant from a reference vector and no longer be differentiable when *r-ordering* takes place, *i.e.,* implying $a = b$ [38, 54]. As such, the approach developed in Chapter 4 incorporates reference colour values locally into the transform for each pixel in a probing SE. By encoding them in the SE, as opposed to being a global reference, non-flat operations are possible and the operation is less sensitive to the anti-symmetry characteristic. Pictorial examples of *m-ordering*, *c-ordering*, and *r-ordering* are given in Section 3.3.2.

Multiple proposals for general colour and multivariate morphology frameworks and colour HMTs in particular exist, each with varying methods for defining morphological operators in multichannel domains. The most common ways morphology has been applied to multivariate spaces is through utilising one of the four ordering relations outlined by Barnett [51].

Multiple linear [55] and non-linear [56–59] filters applied to multi-channel images are reviewed by Comer and Delp in [16]. The authors present two definitions for colour morphology by adapting the techniques used in other multi-channel filters. The first uses an *m-ordering* scheme and the second an *r-ordering* vector-based scheme. The conclusions of this work are that a marginal strategy is effective in certain cases. However, such an approach produces false colour artefacts which are potentially detrimental to further processing steps and the definition of a colour ordering is needed. The vector-based reduced ordering scheme drew from the set of input vectors, avoiding the introduction of new colour vectors.

In [60], the authors use space filling curves to define a vector median filter. Indexing the image using some space filling curve can transform vector data into scalar data which can then be ordered and processed. They then apply the same theory to morphological

operations in [61]. Such space filling curves have an inverse transform and can therefore be used to reconstruct the image after a morphological operation has taken place. Lambert and Chanussot [62] apply similar space filling curves and other bijective transforms to implement a lexicographical (*c-ordering*) set of morphological operators, comparing their algorithm with a marginal approach to good effect.

An alternative method, the so-called majority sorting scheme, is presented in [63] where a vector is ordered based on its number of occurrences within the support of a SE. This approach requires quantisation in order to be viable however, as small variations in image intensity can cause there to be no repetition of colours, and therefore no order can be drawn. This quantisation subsequently introduces both the need to determine the number of colours to be retained as well as the value to which they will be quantised. These choices impact the infimum and supremum and will subsequently effect the outcome of any morphological operations defined using this method.

In [64] the author investigates the combination of *r-ordering* and *c-ordering* using the Mahalanobis colour distance in order to apply morphological operations to colour images. Similarly, a set of operations based on *r-ordering* and *c-ordering* schemes is presented in [38]. The *r-ordering* is based on a reference colour vector and measures the distance to the reference from each other colour vector within the support of the SE. This allows for local supremum and infimum, within the support of the SEs used, to be defined in a similar way to traditional morphology. This approach also allows for various colour space representations to be used. However, neither specifically defines a colour HMT. The ordering scheme presented here is similar to the method employed in the proposed colour HMT detailed in Chapter 4, albeit the reference colours used are defined on a pixel-wise basis by the SE itself, as discussed in Section 4.2.

Similar methods are presented in [54] where a pair of vectors which produce the maximal distance in the support of the SE are defined as the eroded and dilated values. In the case of colour images, in the L*a*b* space, the eroded value typically being the darker (lower L^*) of the two and vice versa for the dilated value. For MSI and HSI data, the eroded and dilated values are defined as the element in the pair with the smaller and larger magnitude respectively. This method is used to define the HMT in [65] where the authors utilise the same maximal distance metric to define a local order on pixels within the support of an SE combined with a POHMT [50] to detect buildings in MSI data.

Weber and Lefèvre [66] define an HMT using a marginal-like strategy, however, rather than use the same SE in each band of the multivariate image as is the case in an *m-ordering* approach,

a specific SE is designed for each band. A hit is defined as the case where every SE matches its respective band in the image. Multiple parameters are introduced for the creation of these extended SEs; the band it is assigned to, a threshold value for noise robustness and a type, whether it is a background or foreground SE. This work was later extended and formalised by Weber and Lefèvre in [67].

In [68], a Vectorial HMT (VHMT) based on a vector *c-ordering* is defined using two colour templates or Structuring Function (SF), a lower SF and an upper SF, where objects that fit between the two are detected. The VHMT is sensitive to noise and therefore a robust VHMT was developed, which utilises rank-order filters. The rank parameters are obtained by defining a maximum allowed number of non-matching pixels in each SE for the object to still be detected. A limitation of using such an approach is that this measure is highly specific to any given application and the object's size and shape. Whereas in the work presented in this thesis, alongside that outlined in [49], a percentage of pixels are used to provide noise robustness and to generalise more appropriately to changes in both object and SE scale.

An extension of the Multiple Objects Matching using Probing (MOMP) HMT [44] to the colour domain is presented in [69, 70]. This introduces the concept of convergence and colour distances within the L*a*b* colour space in order to define colour erosion and dilation. Two points, a convergence and divergence point, are defined and an order can be derived by measuring the distance between each point under the support of an SE to either of these points depending on the operation taking place. This definition allows for the use of non-flat SEs in colour morphology.

In [40] the authors define the Similarity Hit-or-Miss Transform (SHMT) using a combined *r-ordering* and *c-ordering* scheme based on colour similarity. This is a similar approach to that of the proposed multivariate HMT in this thesis, where colour similarity is calculated based on the likeness between the image values and the corresponding values in the SE - allowing for non-flat SEs to be utilised. This SHMT approach has the disadvantage of being restricted to only colour images with three channels due to the definition of similarity functions based on Finlayson's colour image formation model [71]. An order is then defined using a lexicographical cascade where the colour similarity is ordered over the green, red, and blue channels in turn. The authors also present a noise robust variant, which allows for the detection of an object, assuming the lowest similarity value within the SE is above some threshold. Along with the limitation in spectral dimension, a single noisy pixel can cause the minimum similarity value in an object of interest to decrease and in order to detect it, the corresponding threshold must be set so low as to introduce false positive detections. This is discussed in Section 3.3.3 when

comparing the SHMT with the approach proposed in this thesis.

Many of the techniques present in literature, notably those using an *r-ordering* relation to derive a scalar from vector data, require a specific distance, similarity, or disparity function in order to perform the dimensionality reduction operation. There are also proposals for performing the analysis in various colour spaces, each with different optimal distance measures. This naturally leads to the question of what is the "best" colour space and distance function combination? Discussion about which distance measures to use are presented in [72, 73].

In the L*a*b* space, both the Euclidean distance [74] and ΔE measures [54, 69], are used, with the Euclidean distance between two colours being proportional to their difference as perceived by humans [74]. As discussed in [54], the choice of using the L*a*b* colour space instead of RGB for the definition of the colour pseudo-morphological operators is justified by the visually similar results to classical morphology. In addition, the Euclidean distance used in the RGB colour space may produce ambiguous cases when considering the extrema, which does not occur in the L*a*b* space [54].

However for any colour space, the dimensionality of the input data is naturally limited to $N = 3$. For MSI and HSI data, the Euclidean distance measure is a reasonable comparison metric [54]. However, for HSI data in particular, other methods which use the inter-channel correlation can be more optimal [73]. As the aim of the work presented in the first contribution chapter of this thesis concerns the development of a general multivariate HMT, it is apparent that the RGB colour space, due to its similarity in structure to both MSI and HSI data, and the Euclidean distance measure, due to its generalisability and inherent scalability, are suitable for investigation.

2.3 Dimensionality Reduction and Band Selection

HSI data consists of rich spatial and spectral information, which can be useful in several applications. However, the high dimensionality present in hyperspectral images, whilst a powerful resource, often restricts the applicability of such data in certain scenarios, *i.e.,* in low cost or low power tasks. Adjacent bands in hyperspectral images are typically highly correlated and therefore there are often large amounts of redundant information within the image. This can be exploited via compression or DR schemes in order to reduce the number of channels required to present the information within the image.

Principal Component Analysis (PCA) [75] is a popular, classical, method of DR. It seeks to remap highly correlated data into an uncorrelated space using a set of optimal orthogonal basis

vectors, or Principal Components (PCs), calculated from the input data. This can be performed relatively efficiently via an Eigenvalue Decomposition (EVD) where these basis vectors are the eigenvectors of the data, ordered based on the magnitude of the corresponding eigenvalue. Each eigenvalue indicates the proportion of the input data's variance represented by its eigenvector. The original data is transformed using the calculated eigenvectors, where the resultant channels are ordered in terms of their significance. The origins of PCA can be traced back to Pearson [76] amongst others as stated in [75] although, the modern form of PCA was first formally defined in terms of PCs by Hotelling [77] as the Hotelling Transform.

Conventional PCA suffers from three major limitations [78]; 1) calculating the required covariance matrix for high dimensionality datasets is computationally expensive, 2) There is a high computational cost associated with the analysis of large datasets, 3) small local structures can be overlooked in the calculation of PCs. There are multiple variants of the PCA algorithm each with their own advantages and disadvantages when considering the limitations of conventional PCA.

An alternative to PCA is the Maximum/Minimum Noise Fraction (MNF) transformation [79]. PCA can often be affected by noise when calculating PCs and does not guarantee PCs with decreasing quality be produced [79, 80]. MNF is similar to PCA in that it is based on the sorting of eigenvectors based on decreasing Signal-to-Noise Ratio (SNR) as opposed to variance and therefore accounts for the noise present in data [81]. As a result, it is sometimes referred to as the Noise Adjusted Principal Components transform in the literature [82–84].

With both PCA and MNF, as well as many other PCA-like methods, it is necessary to compute the full covariance matrix of an input image. This covariance matrix is a square matrix with sides of length equal to the number of spectral bands in an image. Therefore, for images with high spectral resolution it can be computationally expensive and time-consuming to compute. In order to circumvent this challenge, Folded Principal Component Analysis (FPCA) [85] folds each observation in the data, transforming a vector into a matrix. The covariance for this matrix is then computed for each of the individual observations forming a set of partial covariance matrices. Combining these partial covariance matrices provides the full covariance matrix for the dataset. This, whilst reducing the size of the covariance matrix and simplifying its calculation, also incorporates the correlation between spectral bands into the transform [85].

Similarly, Two-Dimensional PCA (2D-PCA) [86] is based on calculating covariance from an image matrix rather than a matrix of observations from the image. Segmented PCA (Seg-PCA) [87] calculates multiple partial covariance matrices using portions of each observation before combining them. Kernel PCA (KPCA) [88] projects an input dataset into a higher-dimensional

space using the non-linear *kernel trick*. This provides accurate feature detection in low dimensional data, however, in the example of HSI data, where the dimensionality is already large and the aim is to reduce the processing overhead, this is not necessarily preferable. Morphological Principal Component Analysis (MPCA) [89, 90] exploits the structure of data, namely images, within an input set by taking spatial information into account when calculating the covariance matrix.

Independent Component Analysis (ICA) is a linear technique originally devised for Blind Source Separation, separating mixed signals in the so called *cocktail-party problem* [91] into a set of Independent Components (ICs). Unlike PCA, ICA aims to maximise the statistical independence of components as opposed to their variance [92]. Although used for the problem of Blind Source Separation, ICA can be readily applied to hyperspectral classification [93] and DR [94] as each pixel in HSI data can be represented as a weighted mixture of a finite number of Endmembers plus some residual noise. A comparison of PCA, KPCA and ICA for DR in Support Vector Machine (SVM) classifiers is presented in [92]. PCA and MNF are both classified as second order statistics-based transforms which can be insufficient in some applications [95]. ICA preserves higher order statistics which can aid in applications which require characterisation of subtle differences in signature such as classification or in small/rare target detection.

Determining the number of components, or coefficients, to retain in order to produce an optimal representation of the original data is a significant challenge in any DR scheme. In some cases, such as those where the desired number of components is known, it is unnecessary and therefore preferable to avoid calculating every PC as occurs when using an EVD. In these cases, iterative techniques can be used to calculate each successive PC in turn until the required number has been reached. One method to achieve this is by using the Non-linear Iterative Partial Least Squares (NIPALS) algorithm [96].

With PCA, if the desired number of components is unknown, this can commonly be estimated using so-called elbow or scree [97] plots via thresholding or identifying a plateau point. Such graphs show the relative, or sometimes cumulative, proportion of variance contained in each of the PCs, which is equivalent to how much of the data is explained by these retained coefficients. As such, the percentage of variance retained using k components can easily be identified.

With other methods such as ICA no similar graph-based methods exist [94]. Virtual Dimensionality (VD) [98] is a technique for estimating the number of spectrally distinct sources present in an image and is useful when performing Endmember-based DR techniques such as

ICA. Newer developments have seen the notion of VD applied to band selection as well as dimensionality estimation [99].

In addition to classical statistical methods, there is an active area of research concerned with learning an optimal mapping of high dimensional data onto a lower dimensional manifold [100]. The advantages of such a process is that, alongside the forward transform into the manifold, a bi-directional mapping is achieved and data points that are transferred into the lower dimensionality space can be remapped into the original, high-dimensionality, space. This, like other DR methods, also allows for previously unseen data to be transformed into the lower dimensional space. Similarly, autoencoders have been utilised in learning a compressed representation, with the hidden layer being used as a low dimensional feature vector of the input layer [101, 102] and a remapping existing as the output layer.

Where methods such as PCA and its variants aim to project data into a more informative space and are highly effective for DR, they do not preserve the relationship between channels present in an imaged scene, nor do they necessarily provide insight into how such compression was achieved [102]. Band selection however, preserves this inherent structure whilst retaining the most informative channels from the original set for further processing. Band selection is commonly based on maximising some Mutual Information (MI) criteria [103].

Until now, in this thesis, only spectral DR techniques have been covered. Spatial data, especially in remote sensing TD applications where the majority of the scene is often natural vegetation or another highly predictable material, can also be redundant. As a result, there is some potential for spatial DR to be applied using VIs and other similar derived features to identify and exclude known non-target pixels from the analysis.

The NDVI [104] was initially developed to monitor the abundance of vegetation in remote sensing data using observations at two discrete wavelengths in Landsat 1 data; one in the red visible range, the other in the Near-Infrared (NIR). The NDVI gives a measure of abundance of vegetation in each pixel and is one of many vegetation-focussed remote sensing indices reviewed in [105]. In [106], NDVI was used to asses the land cover of Jabalpur City, by thresholding the NDVI values, multiple classes could be determined. It is used similarly in [107, 108] where thresholding and inverting the NDVI values serves as a preliminary mask of non-vegetative target objects.

Variants of NDVI have been developed for use in narrowband and broadband sensors notably the Normalised Difference Vegetation Index (red-edge) ($NDVI_{re}$) [109, 110] which alters the definition of the ratio, using a band in the red-edge portion of the spectrum in place of the red band in base NDVI. Similarly, Red Edge Normalised Difference Vegetation Index

(RENDVI) [111, 112] which uses more specific, narrowband, wavelengths to define the ratio between red-edge and NIR. Other commonly used remote sensing indices are the Normalised Difference Water Index (NDWI) [113, 114] for detecting water bodies in images, or an alternative definition which monitors water in vegetation [115], the Normalised Difference Snow Index (NDSI) [116] which detects snow cover, and the BUI [110] which can be used to separate populated areas from their surroundings. Depending on the imaged scene, each of these indices could be used to isolate otherwise camouflaged, anomalous, objects from their background prior to TD or other identifying analysis.

Remote sensing data can consist of high resolution RGB colour data, radar, multispectral or hyperspectral images. The latter, while providing a great deal of useful information, often at wavelengths beyond the range of human vision, introduces a vast quantity of data which must be handled and processed. DR offers methods for compressing and remapping this high dimensionality data into a reduced, and sometimes more informative, uncorrelated subspace. As hyperspectral images contain high levels of redundancy, they are easily compressed using DR methods. Coupling spectral DR with TD in order to improve detection and classification rates has been discussed in the literature in [84, 117–123], more recently by the authors of [124] as well as some of the author's previous work [125, 126] Appendix C and has been shown to improve the performance of TD and classification algorithms.

2.4 Target Detection

Target Detection from airborne imagery is a major challenge and active area of research within the disciplines of signal and image processing [127, 128]. There have been a wide range of TD algorithms of varying complexities developed over the last few decades [129], ranging from complex mathematical models to those based on more intuitive approaches such as angles or distance between spectra. The most notable difficulties in aerial TD are discussed in [121] and include sensor noise effects, atmospheric attenuation and subsequent correction which can both lead to variabilities in target signature. TD algorithms require *a priori*, or *a posteriori*, information about the target signature to be identified in a Pixel Under Test (PUT). In addition, prior knowledge of any background spectra is sometimes utilised in order to further differentiate the target spectra in an imaged scene. This differs from Anomaly Detection (AD) which requires no prior knowledge about targets or the scene and instead detects outliers based on statistical measures derived from the image data. Various common target and anomaly detectors and their application to various image analysis tasks are discussed here.

In signal processing, a Matched Filter (MF) is used to assess the correlation of a reference signature with an unknown signature to detect the presence of desired features. In TD this is akin to detecting desired spectra in an image and the MF attempts to maximise the response of the known reference in an input signature whilst suppressing the unknown background [118]. Similarly, the Constrained Energy Minimisation (CEM) algorithm [84] seeks to create an optimal Finite Impulse Response (FIR) filter based on the desired target signature and aims to detect the target whilst simultaneously suppressing any background noise. The CEM algorithm is a whitened variant of the MF [130] where the input signatures have not been mean centred. If both the observed spectra and target signature are mean centred the sample correlation is identical to the sample covariance [120, 131].

The Adaptive Coherence/Cosine Estimator (ACE) [132, 133] is a powerful and effective TD algorithm commonly used in remote sensing [123, 131, 134] as well as other signature detection applications [135]. ACE operates under the assumption that an observed sample that corresponds to a target can be modelled as a linear combination of a known target signature and Gaussian noise. By using an estimate of the average background signature and covariance, the coordinate space can be transformed before comparing the target spectra with an observed variable. The process of estimating the covariance of the background, for any input image, is what makes the algorithm *"adaptive"* [136]. Other variants of ACE include Subspace and Simplex ACE [137], which account for the variability in the target material or where a target object is described by a spectral library. ACE has been shown to achieve favourable results in comparisons with other, similar, TD algorithms [120, 121, 124, 131, 138].

In some cases, a direct comparison between image spectra and the reference spectra is sufficient to determine whether the PUT contains target spectra. Two common comparisons are the Spectral Angle Mapper (SAM) [139] and Spectral Information Divergence (SID) [140]. The SAM is a relatively simple method which does not require any image statistics such as covariance or correlation as in ACE or CEM. It calculates the angle between the observed spectral signature and a target signature and is therefore invariant to changes in intensity or illumination. Although a lightweight operation, SAM does not take into account any variations in noise, and small non-linear effects in the data can have an effect on the reliability of the detector [131]. SID is a measure of band-band variability between two vectors, in the case of TD these are the image signature and the reference signature. SID measures the similarity between two signatures by treating them as random variables and measuring the divergence between them, derived from the Kullback-Leibler divergence (KL) information measure [84]. Combining both SAM and SID using *tan* and *sin* functions has been shown to improve upon

both measures [141].

The Orthogonal Subspace Projection (OSP) algorithm [117] differs from those discussed previously as, in addition to knowledge of the target spectrum, *a priori* knowledge of the Endmember spectra present in the background is required. The OSP algorithm first defines a subspace which is orthogonal to the spectra forming the background of an image. It can be assumed that some elements of the target spectra will be orthogonal to this background subspace [142]. The image data is projected onto this subspace which suppresses the background whilst simultaneously improving the desired signature detectability [143]. An MF is then applied to the data within the subspace in order to detect the target signature.

The ability to meet this requirement for knowledge of the exact environment a target will appear in is often unrealistic. An alternative method is to estimate the background using *a posteriori* knowledge gained from the image itself. This can be conducted in two ways; the first concerns estimating the Endmember spectra that form the background using an Endmember extraction algorithm, the second assumes no knowledge and instead the inverse of the sample correlation matrix is used as the projection into the orthogonal subspace. These *a posteriori* measures can be treated as *a priori* knowledge for the purposes of the OSP detector [144, 145]. However, these techniques may result in a sub-optimal or inaccurate detection due to differences between these estimations of the projection and the true subspace projection [145]. This limitation ultimately led to the development of the CEM algorithm in applications where the background is unknown.

As defined in [145], OSP generally requires *a priori* information about the background upon which targets are situated. There are various methods of obtaining the background, including linear transformation such as the Singular Value Decomposition (SVD) or PCA where the most signifiant components which explain the maximum energy of the image, *i.e.,* the background, are retained. Alternatively, the true Endmember spectra of the background can be identified using Endmember extraction algorithms [146]. In [124], the iterative N-FINDR algorithm [147] is used, other Endmember extractors used for both defining the background as well as finding outliers include the Sequential Maximum Angle Convex Cone (SMACC) [148], Pixel Purity Index (PPI) [149], and Vertex Component Analysis (VCA) [150].

Unlike the TD approaches discussed above, AD schemes do not require *a priori* knowledge about a scene or any targets that may be present in it. The Reed-Xiaoli Detector (RXD) [151] is often cited as the benchmark AD method [121] and is equivalent to the Mahalanobis distance between the observed spectral signature and the average signature of the background. Alternatively, as presented in [152], the RXD detects targets in an image that have small

energies represented by small eigenvalues, *i.e.,* the RXD can be considered the mathematical inverse of PCA [152, 153]. The background signature can be estimated using the entire image, Global-RXD, or from a region local to each pixel, Local-RXD. In this thesis RXD refers to the use of Global-RXD. As the RXD is a measure of distance between an observation and the average signature, higher values will correspond to outliers that may be of interest and often highlights targets, although whether or not they are detected is determined by their difference from their background. A review of many of the AD methods common in literature is presented in [154].

In [119], TD algorithms are applied to subspaces generated by retaining varying numbers of PCs and it is observed that they remain consistent and robust detecting targets in a subspace. In [155], the authors use the OSP detector combined with the PCA and MNF DR schemes for both noise removal and compression. This method alters the way in which the OSP operator is implemented and as such requires only *a priori* information about the target signature. In [124], the authors explore using DR with various TD algorithms and investigating the performance of the detection within the low dimensionality subspace. In previous work, they also investigated the reconstruction of the original data from the DR coefficients [156]. Coupling spectral DR with TD in order to improve detection and classification rates has been discussed in the literature [84, 120–123, 138] and has been shown to improve the performance of TD and classification algorithms.

Other methods such as the RXD are sensitive to kurtosis [154] and therefore a DR algorithm that exploits this fact, amplifying the level of outliers, can improve their detection. However, many DR methods seek to condense information onto few optimal basis vectors, in the process removing outliers and other statistically insignificant measures [153]. As a result they are less often utilised with AD schemes. However, in [154], a subspace RXD is discussed where the first, and statistically least outlying, PCs are discarded.

2.5 Compressed Sensing

High dimensionality data such as multispectral and hyperspectral images offers a vast quantity of information about a scene, however, this large volume of data requires both efficient management and processing techniques. Hyperspectral images in particular often exhibit large amounts of redundancy, therefore, compression and DR techniques can frequently be used to good effect. However such techniques require the image to be captured in its entirety.

In image compression standards, it is known that images can be compressed in a fixed

way allowing for the efficient storage and transfer of images. Any signal of length N can be represented in terms of a basis of $N \times 1$ vectors. The signal is called K-sparse if it can be represented as a linear combination of only K basis vectors, *i.e.,* if there are K non-zero components in some vector of coefficients [157]. If only K components are non-zero the signal is truly sparse. However, more likely is that there are K significant components, and $N - K$ small coefficients that can be discarded [158]. In such cases the signal is said to be compressible and can be well approximated from sparse representations [159]. However, the need to capture the full representation of an image prior to wasteful transfer coding-based compression remains [160]. Whereas DR and compression schemes seek to exploit statistical redundancy in high-dimensionality data, CS techniques seek to directly acquire a compressed signal representation of the image [159]. This can be achieved using various structured sensing techniques [158].

CS has been a rapidly advancing research topic in multiple areas of signal processing since Candès, Romberg and Tao [161–163] and Donoho [164] demonstrated that a signal can be reconstructed from a sparse set of linear, non-adaptive measurements. CS approaches commonly involve two stages: the first, capturing a compressed representation of a scene; and second, the reconstruction of an image from this compressed representation prior to any subsequent processing and analysis [160]. This reconstruction is often achieved using convex optimisation algorithms such as Gradient Projection for Sparse Reconstruction (GPSR) [165], Two-step Iterative Shrinkage Thresholding (TwIST) [166], and Alternating Direction Method of Multipliers (ADMM) [167]. Recently, deep learning techniques have also been used to learn the reconstruction of CS measurements [168–170].

The same sparsity assumptions that are made for signals can be applied to image and video data, where each pixel is represented by a vector [158]. In particular, the sparse component of HSI data can be separated from the background using Low-Rank and Sparse Matrix Decomposition (LRaSMD) [171]. LRaSMD exploits the assumption that HSI data consists of a finite number of Endmember present in the image weighted by a set of abundances with a sum-to-one constraint. As such, they can be represented as a low-rank matrix plus a sparse component and random noise [172].

In TD and AD applications, the targets of interest populate this sparse part of the image and by separating it from the data as a whole, the target visibility can be increased. LRaSMD and other similar approaches are used frequently in the literature in remote sensing TD applications [172–174] as well as other applications such as medical imaging [175]. LRaSMD can be applied to video data in order to extract sparse objects from their background [176]. Unlike CSI

techniques, the entire image or video scene must be captured prior to the extraction of the sparse component and any subsequent processing to exploit sparsity using LRaSMD.

One well known example of CSI hardware is the Single Pixel Camera (SPC), developed by Duarte et al. at Rice University [177] soon after the initial development of CS theory. The SPC uses a Digital Micrometer Device (DMD) to reflect light onto a photodiode based on some pattern [178]. Alternatively, structured illumination can project known patterns onto a scene, with the response being recorded [179]. Commonly, orthogonal patterns such as the Hadamard basis are used [178]. The response at a single photo-detector characterises the scene as it is encoded or illuminated by each pattern and by illuminating the scene with thousands of patterns a set of measurements is obtained. By repeating over multiple successive patterns, sufficient measurements can be taken to reconstruct an image by weighting each pattern and associated response [180] in a similar way to other classical compression methods such as the Discrete Cosine Transform (DCT). SPC-based methods have been applied to various image analysis tasks, usually as a low-cost, enabling technology [178, 179]. With adaptation, the SPC is scalable and has been applied to both video [181] and HSI data [177].

In order to obtain an acceptable quality reconstruction, SPC architectures often use thousands of illuminations [181], this requires long exposure times when compared to architectures with 2D sensing capabilities [182]. In order to alleviate this, multi-aperture and side-aperture systems were developed, these help guide the reconstruction process reducing the number of necessary samples [183].

Another system for CS in hardware commonly used for CSI is Coded Aperture Snapshot Spectral Imaging. Where the SPC captures a set of scalar measurements based on the illumination of a scene, CASSI architectures capture a coded version of an image that has been passed through a structured pattern, or Coded Aperture, before being spatially multiplexed through various optical elements [184] and finally arriving incident on a sensor. This process is described in detail in Chapter 3. Similar to the multiple illuminations of the SPC, multiple images of a scene are taken each with a different CA, and the result is a set of 2D spatio-spectral coded representations of the imaged scene. These can then be reconstructed from the compressed domain into the image domain via convex optimisation. CASSI architectures vary in terms of the optical elements used within [185] and have been used in various imaging areas such as super-resolution [186, 187], image classification [188, 189], remote sensing [190], and target detection [190, 191].

Critical to any CASSI system is the design of appropriate CAs [192, 193]. Initially, CAs consisted of 2D binary patterns which allowed for incoming light to be either transmitted or

blocked [184] at a given spatial location. There have been many investigations into generating ensemble CAs which optimally sample an image so as to obtain a high-quality reconstruction [184, 194–196] over multiple, successive, snapshots. Blue-noise sampling in particular provides a highly optimal spacing between samples as well as suppressing low frequency components of noise [193, 197]. Additionally, with the recent advent of deep learning, approaches to learn optimal sets of CAs have been investigated [198].

Recently there has been intensive research into improving the performance of CASSI-based systems by implementing colour and spectral bandpass filters into the CAs [199–202]. This allows for a greater spectral selectivity which improves applications where including *a priori* information about a task is important such as TD [185, 203].

Other snapshot compressive spectral imaging techniques include the Spatial-Spectral encoded Compressive Spectral Imager (SSCSI) [204], the Snapshot Coloured Compressive Spectral Imaging (SCCSI) [196], the Dual-Coded hyper-Spectral Imager (DCSI) [205], amongst others [185]. Many of these snapshot techniques are limited by the resolution of the DMD applying the spatial coding. As a result, there have been attempts to use high-resolution side information in order to upscale and guide the reconstruction [183, 195, 206, 207].

Other imaging techniques have been developed and used to capture sparse HSI data and perform TD in the reconstructed domain. One very different technology from snapshot imagers is the Compressive Sensing Miniature Ultra-Spectral Imaging (CS-MUSI) [208] architecture which, by varying voltage applied across a liquid crystal element, uses the properties of birefringence with the refractive index changing based on the voltage applied. This allows for images at different wavelengths to be captured as a function of the voltage and has been applied to TD applications [209] as well as general HSI data capture [208]. Similar to the work proposed in this thesis the CS-MUSI can perform target detection in the compressive domain without the need for reconstruction [209, 210].

Multiple studies have shown that reconstructing scenes from CASSI measurements can provide accurate results in classification tasks [185, 188, 189]. However, the reconstruction of an image from encoded measurements can be a complex, and time consuming, optimisation task. Other proposed methods seek to perform analysis of MSI and HSI without the need for reconstruction, instead extracting information from the compressed measurements. The CS-MUSI camera [209] has been used in TD [210] and classification [211] tasks directly from compressed measurements. Similarly, an approach using a Convolutional Neural Network (CNN) is used to classify the compressed representations of a scene captured using a CASSI sensor [212]. A similar approach for performing TD using a CASSI system is of interest as

CS techniques in general reduce the amount of memory and processing power required to understand a scene, and if sufficient information is retained in the compressive measurements, the need for reconstruction prior to performing TD can be removed.

2.6 Summary

In this chapter, the relevant literature pertaining to each of the research areas considered in this thesis were introduced. This includes the original definitions of Mathematical Morphology and its extension to greyscale images and beyond. Various Dimensionality Reduction and Target Detection schemes, and how they can be used in hyperspectral image analysis have been presented along with methods of applying Compressed Sensing to hyperspectral image capture. The three main contribution areas presented in Chapters 4 to 6, are based on the topics discussed in this chapter with the unifying theme of how they can be applied to remotely sensed hyperspectral image data in order to facilitate accurate and efficient object detection and classification. The mathematical fundamentals of each of the techniques discussed here, as well as some general definitions, are presented in Chapter 3.

Chapter 3

Theoretical Background

3.1 Introduction

This chapter introduces the theoretical and mathematical backgrounds of each of the various techniques utilised in the novel contributions of this thesis. Firstly, in Section 3.2 the mathematical fundamentals of scalar morphology are presented. In Section 3.3, an outline of the multivariate ordering schemes discussed in Section 2.2.2 are shown, as well as the definition of the SHMT, a technique compared with the contribution developed in Chapter 4. Section 3.4 gives an overview of Hyperspectral Imaging while Sections 3.5 and 3.6 describe the Dimensionality Reduction and Target Detection algorithms used in this thesis respectively. The mathematical foundations of Compressed Sensing and descriptions of CASSI architectures are presented in Section 3.7. While not a focus of this thesis, deep learning methods for object detection are discussed in Section 3.8. Finally, Section 3.9 provides a summary of the topics covered with respect to the aims of this work.

3.2 Scalar Mathematical Morphology

Mathematical Morphology is well defined for single channel binary and greyscale images. In this section, the mathematical foundations of MM are covered and the fundamental operations are derived.

3.2.1 Binary Image Formation

Pixels in a binary image can take the values 0 or 1, being either "off" or "on", and belong to the Foreground (FG) or the Background (BG) sets. In this thesis, "on" pixels take the value of 1 and represent the foreground of the image and "off" pixels take the value of 0 and represent the background of the image, *i.e.,* having a dark support [27]. An example of a binary image can be seen in Figure 3.1.

Figure 3.1: Example of a binary image.

A binary image, f, can be considered as a mapping of the definition domain, or image plane, of f, $\mathcal{D}_f$, where $\mathcal{D}_f$ is a subset of the two-dimensional Cartesian space $\mathbb{Z}^2$, into the set of values {0,1} [27], *i.e.,*

$$f : \mathcal{D}_f \subset \mathbb{Z}^2 \rightarrow \{0, 1\} \tag{3.1}$$

where $\forall x \in \mathcal{D}_f$, $f(x)$ is either 0 or 1. With this mapping of an image into two sets, MM in binary images can be implemented using techniques from set theory.

3.2.2 Greyscale Image Formation

Pixels in a greyscale image can take on a number of values dependent on the bit-depth of the image. Rather than a mapping into $\{0, 1\}$, as for a binary image, a greyscale image in $\mathcal{D}_f$ is mapped into the bounded set $\{0, 1, ..., 2^{n-1}\}$ [27], where n is the number of bits used to represent each pixel. This mapping can be denoted as:

$$f : \mathcal{D}_f \subset \mathbb{Z}^2 \rightarrow \{0, 1, ..., t_{\max}\} \tag{3.2}$$

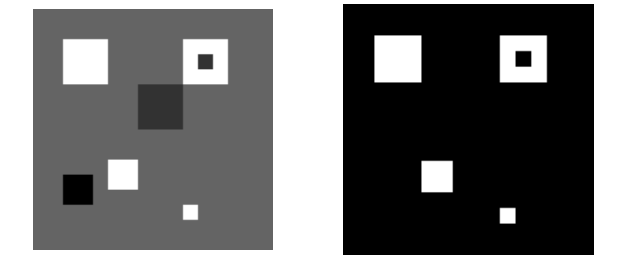
where $\forall x \mathcal{D}_f$, $f(x)$ belongs to $\{0, 1, ..., t_{\max}\}$ where $t_{\max}$ represents the maximum value *e.g.,* $t_{\max} = 2^n - 1$ for an n-bit image.

Greyscale images can be interpreted as a topography, where the intensity of the pixel corresponds with its resultant height within the surface, *i.e.,* bright pixels form peaks and dark pixels form troughs. This surface forms the graph of the image and is the set of points belonging to the image plane [27] with the subgraph comprised of the pixels lying below the graph.

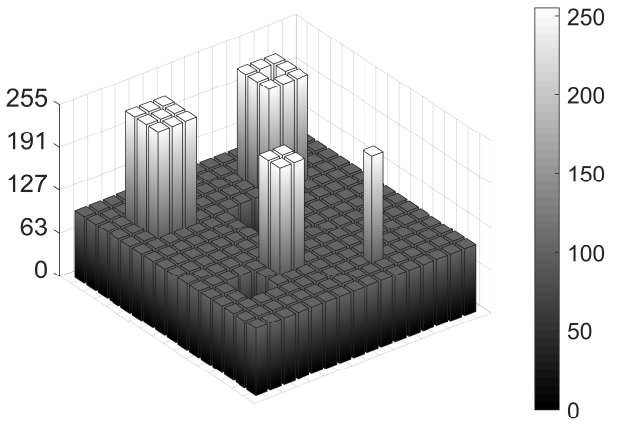
Additionally, each subgraph, or grey-level, t, within the image, with the exception of $t = 0$, can be viewed as a single binary image. This represents a slice from the topography and is

denoted as the cross section of the image, $\mathrm{CS}_t(\boldsymbol{X})$, where pixels are "on" if their value is greater than or equal to the current threshold, t. The topography is obtained by stacking these cross-sectional binary images together.

An example of a greyscale image, $\boldsymbol{X}$, is shown in Figure 3.2a with the corresponding topography in Figure 3.2c. The binary cross section of $\boldsymbol{X}$ at $t = 120$, *i.e.,* $\mathrm{CS}_{120}(\boldsymbol{X})$, is shown in Figure 3.2b.



(a) (b)



(c)

Figure 3.2: Representations of a greyscale image: **a)** Intensity image, **b)** Cross section of a) at $t = 120$, **c)** Image as a topographical surface.

3.2.3 Erosion and Dilation

The fundamental operators of MM are erosion and dilation. The erosion by an SE, $\boldsymbol{S}$, $\varepsilon_{\boldsymbol{S}}(\boldsymbol{X})$, is the locus of all points in a binary image, $x \in \boldsymbol{X}$, where the SE translated to point x, $\boldsymbol{S}_x$, is a subset of $\boldsymbol{X}$:

$$\varepsilon_{\boldsymbol{S}}(\boldsymbol{X}) = \{x | \boldsymbol{S}_x \subseteq \boldsymbol{X}\} \tag{3.3}$$

In the case of a binary image, this is equivalent to all points in the image where a translated SE matches, or fits inside, the foreground. This can be written in terms of the intersections of these translations within the SE:

$$\varepsilon_{\boldsymbol{S}}(\boldsymbol{X}) = \bigcap_{s \in \boldsymbol{S}} \boldsymbol{X}_{-s} \tag{3.4}$$

Dilation is the dual operation of erosion, *i.e.,* a dilation on the background of an image is equivalent to an erosion of the foreground [27]. The dilation of an image, $\boldsymbol{X}$, by an SE, $\boldsymbol{S}$, $\delta_{\boldsymbol{S}}(\boldsymbol{X})$, is the locus of all points where the intersection of $\boldsymbol{S}$ and $\boldsymbol{X}$ is non-empty:

$$\delta_{\boldsymbol{S}}(\boldsymbol{X}) = \{x | \boldsymbol{S}_x \cap \boldsymbol{X} \neq \varnothing\} \tag{3.5}$$

In the case of a binary image, this is equivalent to all points in the image where a translated SE touches, or intersects with, the foreground. This, much like Equation (3.3), can be written in terms of the translations, specifically the unions of these translations within the SE:

$$\delta_{\boldsymbol{S}}(\boldsymbol{X}) = \bigcup_{s \in \boldsymbol{S}} \boldsymbol{X}_{-s} \tag{3.6}$$

MM operations rely on the definition of a total, or at least partial, order and can be executed by exploiting this underlying ordering relation. If an ordering relation exists, or can be introduced, within the data being analysed, morphological operators can be readily applied to images in the region defined by some SE [35, 37, 213]. In the binary case, as with single channel images more generally, this ordering is trivial. Standard erosions can be computed by employing a minimum filter and, conversely, dilations are computed using maximum filters. This gives the following definition of an erosion at a pixel, $x \in \boldsymbol{X}$, where the eroded value is the minimum value in the region defined by the SE, $\boldsymbol{S}$:

$$[\varepsilon_{\boldsymbol{S}}(\boldsymbol{X})](x) = \min_{s \in \boldsymbol{S}} \boldsymbol{X}(x + s) \tag{3.7}$$

Similarly, the dilated value at point x can be defined as the maximum pixel value within the support of the SE, $\boldsymbol{S}$, when it is translated to x:

$$[\delta_{\boldsymbol{S}}(\boldsymbol{X})](x) = \max_{s \in \boldsymbol{S}} \boldsymbol{X}(x + s) \tag{3.8}$$

The images produced as result of performing an erosion and dilation on a binary image, Figure 3.3a, using a 3×3 square SE can be seen in Figures 3.3b and 3.3c respectively.

(a) (b) (c)

Figure 3.3: Example of binary erosion and dilation: **a)** Original binary image **b)** Image after erosion by a square, 3×3 , SE **c)** Image after dilation by a square, 3×3 , SE.

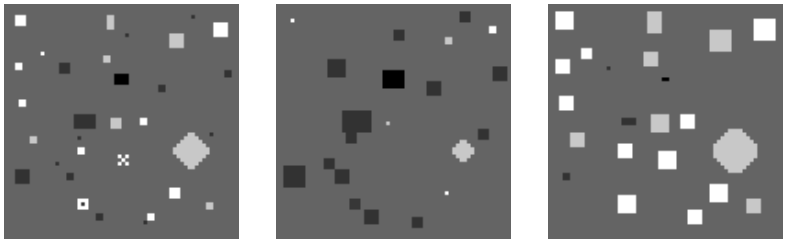
Generally, an erosion has the effect of shrinking objects in the foreground based on the SE used, whereas dilation has the opposite effect and expands them. Alternatively, this can be considered as expanding and shrinking the background respectively.

The translation-based definitions for erosion, Equation (3.4), and dilation, Equation (3.6), on binary images, can be extended to greyscale images by defining the erosion, $\varepsilon_{\boldsymbol{S}}(\boldsymbol{X})$, and dilation, $\delta_{\boldsymbol{S}}(\boldsymbol{X})$, as the point-wise minimum and maximum of the translations of an image, $\boldsymbol{X}$, by the vectors, -s, of the SE, $\boldsymbol{S}$, respectively;

$$\varepsilon_{\boldsymbol{S}}(\boldsymbol{X}) = \bigwedge_{s \in \boldsymbol{S}} \boldsymbol{X}_{\text{-s}}, \tag{3.9}$$

$$\delta_{\boldsymbol{S}}(\boldsymbol{X}) = \bigvee_{s \in \boldsymbol{S}} \boldsymbol{X}_{\text{-s}}. \tag{3.10}$$

Erosion and dilation of greyscale images is easily implemented using the same concept as in binary images, with minimum, Equation (3.7), and maximum, Equation (3.8), filters respectively. Examples of the erosion and dilation of a greyscale image, Figure 3.4a, using a 3×3 square SE are shown in Figures 3.4b and 3.4c respectively.



(a) (b) (c)

Figure 3.4: Example of greyscale erosion and dilation: **a)** Original greyscale image **b)** Image after erosion by a square, 3×3, SE **c)** Image after dilation by a square, 3×3, SE

3.2.4 Opening and Closing

Both erosion and dilation are non-linear operations as they depend on the underlying structure of the image rather than being a simple linear transformation of the pixel values. Additionally, no perfect inverse erosion or dilation exists as the underlying structure of the image is altered by the original operation. As a result, following a basic morphological operation, recovering the original image data is not possible.

For example, in an erosion, objects in the foreground are often completely removed, as shown in Figures 3.3b and 3.4b, and applying a dilation often removes small geometric features from an image, Figures 3.3c and 3.4c. Whilst this behaviour is beneficial in some cases, *e.g.,* for the selective removal of noisy elements which do not fit the designed SE criteria, erosion and dilation are inherently destructive to the remaining foreground and background objects. As a result, techniques for recovering a maximal amount of information about an image following an erosion or dilation were developed resulting in the definitions of morphological opening and closing respectively [27].

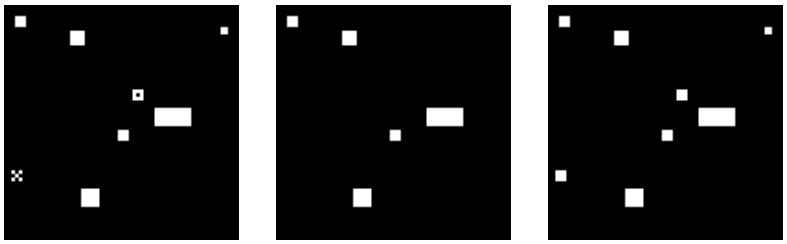
The opening, γ, of an image, $\boldsymbol{X}$, by an SE, $\boldsymbol{S}$, is denoted as $\gamma_{\boldsymbol{S}}(\boldsymbol{X})$ and is defined as the erosion of $\boldsymbol{X}$ by $\boldsymbol{S}$, followed by the dilation of the resulting image with the reflected SE, $\check{\boldsymbol{S}}$, where the reflected, or transposed, SE is mirrored on its horizontal and vertical axes about its origin:

$$\gamma_{\boldsymbol{S}}(\boldsymbol{X}) = \delta_{\check{\boldsymbol{S}}}(\varepsilon_{\boldsymbol{S}}(\boldsymbol{X})) \tag{3.11}$$

Conversely, the dual operation of the opening, the closing, φ, of an image, $\boldsymbol{X}$, by an SE, $\boldsymbol{S}$, is denoted as $\varphi_{\boldsymbol{S}}(\boldsymbol{X})$ and aims to recover the original structure in an image following a dilation. A closing is defined as the dilation of $\boldsymbol{X}$ by $\boldsymbol{S}$, followed by the erosion of the resulting image with the reflected SE, $\check{\boldsymbol{S}}$:

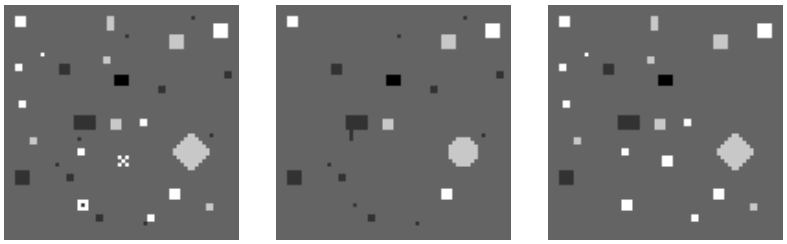
$$\varphi_{\boldsymbol{S}}(\boldsymbol{X}) = \varepsilon_{\check{\boldsymbol{S}}}(\delta_{\boldsymbol{S}}(\boldsymbol{X})) \tag{3.12}$$

Examples of both opening and closing are shown in Figure 3.5 for a binary image and Figure 3.6 for a greyscale image.



(a) (b) (c)

Figure 3.5: Example of binary opening and closing: **a)** Original binary image **b)** Image after opening by a square, 3×3, SE **c)** Image after closing by a square, 3×3, SE



(a) (b) (c)

Figure 3.6: Example of greyscale opening and closing: **a)** Original greyscale image **b)** Image after opening by a square, 3×3 , SE **c)** Image after closing by a square, 3×3 , SE

An opening produces an approximate reconstruction of an image following an erosion and can be used to remove holes, or noisy pixels, in the background of an image. In an opening operation, areas in the foreground of the image where the SE cannot fit are removed in the initial erosion despite the subsequent dilation. Areas in the foreground where the SE can fit, and were shrunk in the initial erosion, are largely restored. Conversely, a closing operation produces an approximate reconstruction of an image following a dilation, where holes in the foreground, *i.e.,* areas which are smaller than the translated SE, can be smoothed in the initial dilation and are largely unrecoverable despite the subsequent erosion.

3.2.5 The Hit-or-Miss Transform

The HMT is a useful tool in MM and is often used for detecting objects based on their size and shape. In common with many other MM operators, the definition of the HMT was originally

restricted to binary images [14]. Multiple proposals for greyscale extensions followed, which are reviewed in [46], and a unified formulation now exists [214].

The HMT differs from both erosion and dilation in the fact that it requires two non-intersecting SEs. The foreground SE, $\boldsymbol{S}_{FG}$, is designed to fit the object under study in an image, $\boldsymbol{X}$, and the background SE, $\boldsymbol{S}_{BG}$, is designed to fit the background of the object or the complement, $\boldsymbol{X}^c$, of image, $\boldsymbol{X}$. These two SEs can be collated into one composite SE, $\boldsymbol{S}$, which is defined as $\boldsymbol{S} = [\boldsymbol{S}_{FG}, \boldsymbol{S}_{BG}]$, where $\boldsymbol{S}_{FG} \cap \boldsymbol{S}_{BG} = \varnothing$. The HMT of $\boldsymbol{X}$ using a composite SE, $\boldsymbol{S}$ is defined as:

$$\mathrm{HMT}_{\boldsymbol{S}}(\boldsymbol{X}) = \{x | (\boldsymbol{S}_{FG})_x \subseteq \boldsymbol{X}, (\boldsymbol{S}_{BG})_x \subseteq \boldsymbol{X}^c\} \tag{3.13}$$

An image pixel $x \in \boldsymbol{X}$ is only output in the foreground set following a HMT when both of the conditions in Equation (3.13) are met. As a result, the notation in Equation (3.13) can be thought of as the intersection of two separate erosions. One where $\boldsymbol{S}_{FG}$ is a subset of the object under observation and $\boldsymbol{S}_{BG}$ is a subset of the foreground of the complement image. Using the definition of an erosion from Equation (3.3), an alternative representation of the binary HMT can be defined:

$$\mathrm{HMT}_{\boldsymbol{S}}(\boldsymbol{X}) = \varepsilon_{\boldsymbol{S}_{\boldsymbol{FG}}}(\boldsymbol{X}) \cap \varepsilon_{\boldsymbol{S}_{BG}}(\boldsymbol{X}^c) \tag{3.14}$$

An example of the HMT on a synthetic binary image is shown in Figure 3.7. A 3×3 white square is used as the foreground SE and a 5×5 square with a 3×3 hollow square removed from it as the background SE. This configuration is designed to detect 3×3 white squares in the foreground on a black background.

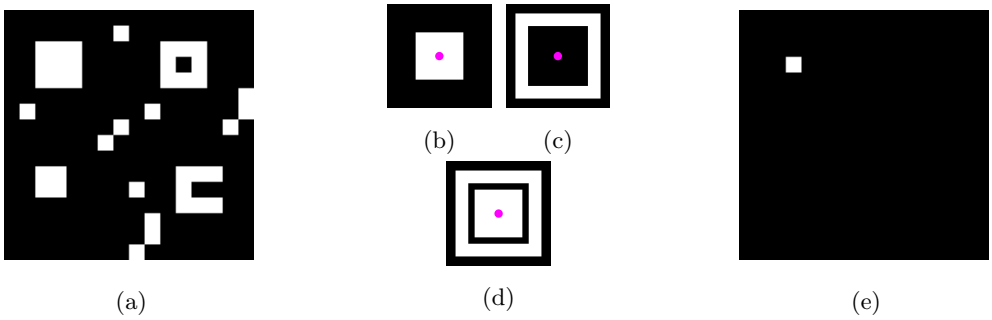
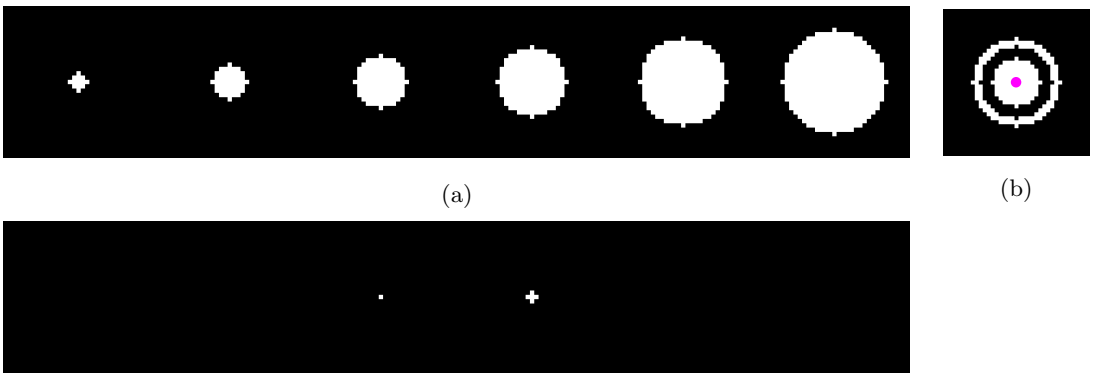


Figure 3.7: Example of a binary HMT: **a)** Binary test image made up of multiple sizes of shapes. **b)** Foreground SE comprised of a 3×3 white square. **c)** Background SE comprised of a 5×5 white square with a 3×3 hole hollowed from its centre. **d)** Binary composite SE. **e)** Output of the HMT of a) using d).

In Figure 3.7 the SEs are designed to fit inside and around 3×3 squares exactly, with no buffer between the foreground SE, Figure 3.7b, and the background SE, Figure 3.7c. The

composite SE, Figure 3.7d, is presented with a barrier for ease of visualisation but in reality would be indistinguishable from a 5×5 foreground square. This convention is continued unless explicitly stated.

SEs can also be designed to detect shapes of multiple sizes by introducing a gap between the foreground and background SEs. Figure 3.8a contains discs of radius $r \in \{2, 4, 6, 8, 10, 12\}$ presented on a blank background. A composite SE, consisting of a circular foreground SE with $r = 6$ and an annular background SE with $r_{\text{inner}} = 9$ and $r_{\text{outer}} = 11$, was designed to detect a subset of these discs and is shown in Figure 3.8b with the result of performing the HMT shown in Figure 3.8c.



(a) (b)

(c)

Figure 3.8: Example of a binary HMT designed for multi-scale objects: **a)** Binary test image consisting of multiple sizes of shapes. **b)** Binary composite SE. **c)** Output of the HMT of a) using b).

By adding separation between the foreground and background SEs, both the discs with $r = 6$ and $r = 8$ were detected. As the foreground SE exactly matches the foreground of the image in Figure 3.8a when translated to the centre of the disc where $r = 6$, there is a single positive detection at the corresponding pixel in Figure 3.8c. However, when considering the disc with $r = 8$, there are multiple points where the foreground SE fits inside the foreground of the image whilst the background SE simultaneously fits around the foreground object. As a result, there are multiple pixels corresponding to this object in the set of detections in Figure 3.8c.

Each of the other objects in Figure 3.8a are not detected as; for the discs with $r = 2$ and $r = 4$, the foreground SE does not fit within the objects, and for the discs with $r = 10$ and $r = 12$ the foreground SE does fit within the objects, however, the background SE also intersects with them, preventing their detection.

Equation (3.14) could be extended to greyscale images but, as discussed in Section 2.2.1, this requires taking the complement of a greyscale image, the result of which would ultimately be

dependent on the value used for complementing the image. In addition to this, the HMT is not an increasing transform, this is to say that if $x \leq y$ then $\text{HMT}(x) \not\leq \text{HMT}(y)$. Therefore stacking the HMT of each cross-section, $\text{CS}_t(\boldsymbol{X})$, and stacking them would not provide a valid solution, both of these factors mean its extension to greyscale is not a straightforward endeavour [27].

Soille defines two greyscale HMTs in [27]: the Unconstrained Hit-or-Miss Transform (UHMT) and the Constrained Hit-or-Miss Transform (CHMT). These are the greyscale HMTs used as examples in this work, as discussed in Section 2.2.1 there have been various other extensions proposed. The UHMT of an image, $\boldsymbol{X}$, is defined as the number of cross-sections, $\text{CS}_t(\boldsymbol{X})$, where the foreground SE, $\boldsymbol{S}_{FG}$, matches $\text{CS}_t(\boldsymbol{X})$ and the background SE, $\boldsymbol{S}_{BG}$, matches its complement, $\text{CS}_t^c(\boldsymbol{X})$, as shown in Equation (3.15) below:

$$[\text{UHMT}_{\boldsymbol{S}}(\boldsymbol{X})](x) = \text{card}\{t | (\boldsymbol{S}_{FG})_x \subseteq \text{CS}_t(\boldsymbol{X}), (\boldsymbol{S}_{BG})_x \subseteq \text{CS}_t^c(\boldsymbol{X})\} \tag{3.15}$$

This definition of the UHMT is equivalent to the cardinality of intersections between the two intervals of levels $[0, t_{\text{FG}}]$ and $[t_{\text{BG}}, t_{\max}]$:

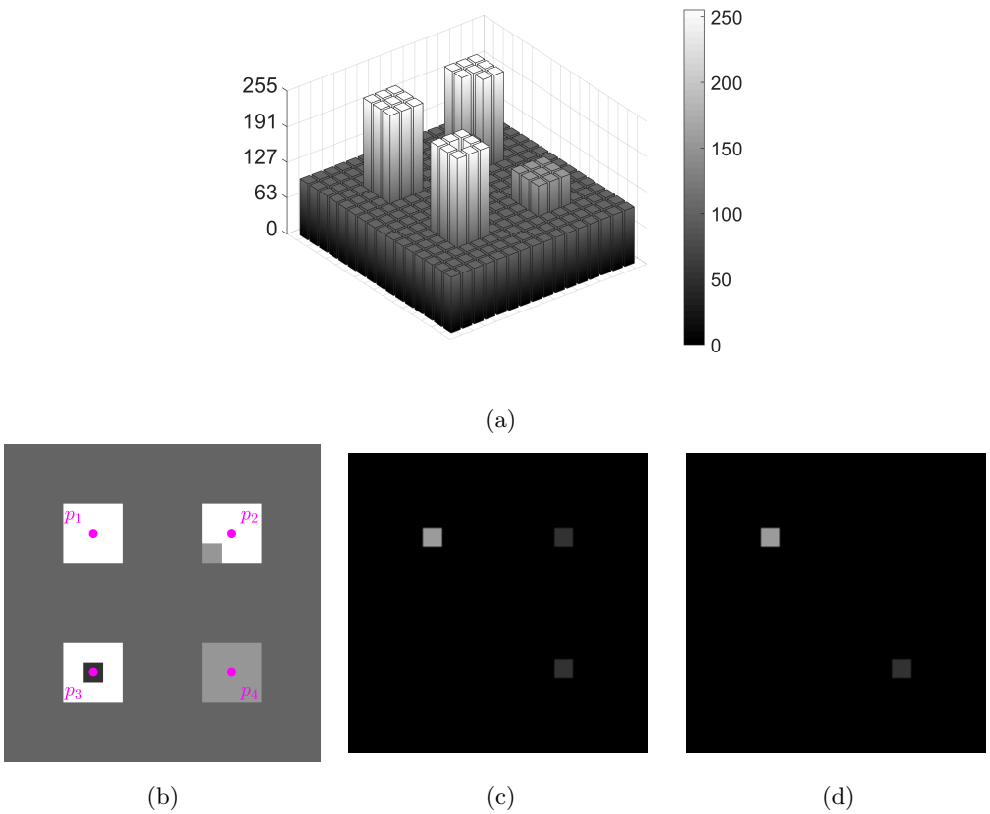
$$\begin{aligned}[\text{UHMT}_{\boldsymbol{S}}(\boldsymbol{X})](x) &= \text{card}([0, t_{\text{FG}}] \cap [t_{\text{BG}}, t_{\max}]) \\ &= \begin{cases} t_{\text{FG}} - t_{\text{BG}} + 1, & \text{if } t_{\text{FG}} \geq t_{\text{BG}} \\ 0, & \text{otherwise} \end{cases}\end{aligned} \tag{3.16}$$

where, $t_{FG} = \max\{t | (\boldsymbol{S}_{FG})_x \subseteq \text{CS}_t(\boldsymbol{X})\}$ and $t_{\text{BG}} = \min\{t | (\boldsymbol{S}_{BG})_x \subseteq \text{CS}_t^c(\boldsymbol{X})\}$. Additionally, it can be noted that $\forall x \in \boldsymbol{X}$, the value of t_{FG}, when $\boldsymbol{S}_{FG}$ is translated to x, is equal to $[\varepsilon_{\boldsymbol{S}_{FG}}](x)$. Similarly, t_{BG}, when $\boldsymbol{S}_{BG}$ is translated to x, is equal to $1 + [\delta_{\boldsymbol{S}_{BG}}](x)$ [27]. As a result, using the definitions of greyscale erosion and dilation in Equations (3.7) and (3.8), the UHMT can be defined as:

$$[\text{UHMT}_{\boldsymbol{S}}(\boldsymbol{X})](x) = \begin{cases} [\varepsilon_{\boldsymbol{S}_{FG}}](x) - [\delta_{\boldsymbol{S}_{BG}}](x), & \text{if } [\delta_{\boldsymbol{S}_{BG}}](x) < [\varepsilon_{\boldsymbol{S}_{FG}}](x) \\ 0, & \text{otherwise} \end{cases} \tag{3.17}$$

This alternative definition of the UHMT has the benefit of removing the need to take the complement of a greyscale image, exploiting the duality between the erosion and dilation operations.

The CHMT is similar, however, with the additional constraint that the t level of the pixel containing the origin point of the SE must match the value of the image graph. This is equivalent



(a)

(b) (c) (d)

Figure 3.9: Comparison of Soille's greyscale HMTs: **a)** Greyscale test image as a surface **b)** Greyscale test image **c)** The unconstrained HMT result on the image in Figure 3.9b **d)** The constrained HMT result on the image in Figure 3.9b.

to the SE fitting inside, or around, the topographical surface of a greyscale image for every level, t, when the origin is contained in $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ respectively. The CHMT is defined in Equation (3.18):

$$[\text{CHMT}_{\boldsymbol{S}}(\boldsymbol{X})](x) = \begin{cases} \boldsymbol{X}(x) - [\delta_{\boldsymbol{S}_{BG}}](x), & \text{if } \boldsymbol{X}(x) = [\varepsilon_{\boldsymbol{S}_{FG}}](x) \text{ and } [\delta_{\boldsymbol{S}_{BG}}](x) < \boldsymbol{X}(x) \\ [\varepsilon_{\boldsymbol{S}_{FG}}](x) - \boldsymbol{X}(x), & \text{if } \boldsymbol{X}(x) = [\delta_{\boldsymbol{S}_{BG}}](x) \text{ and } [\varepsilon_{\boldsymbol{S}_{FG}}](x) > \boldsymbol{X}(x) \\ 0, & \text{otherwise} \end{cases} \tag{3.18}$$

An example of the differences between the UHMT and the CHMT on the test image from Figure 3.2 can be seen in Figure 3.9. The test image from Figure 3.2, denoted as $\boldsymbol{X}$, contains four square objects on a monochromatic background with intensity of $t = 100$:

– The top left square is a white 3×3 square with intensity $t = 255$ making it above the level of the background. As a result, the erosion of the foreground is greater than the

dilation of the background and this object is detected using the UHMT with a value of $[\text{UHMT}_{\boldsymbol{S}}(\boldsymbol{X})](p_1) = 155$ in Figure 3.9c. Additionally, as the eroded value at the origin is equal to the value of image, this object is also detected by the CHMT in Figure 3.9d, where $[\text{CHMT}_{\boldsymbol{S}}(\boldsymbol{X})](p_1) = 155$.

- The top right square is similar in construction to the top left square; however, has a single pixel with reduced intensity, $t = 150$. As this is an intensity greater than that of the background, the erosion of the foreground still produces a value greater than the dilation of the background and the square is detected by the UHMT with a value of $[\text{UHMT}_{\boldsymbol{S}}(\boldsymbol{X})](p_2) = 50$. However, as the eroded value is non-equal to the value of the image graph, this square is not detected using the CHMT. It should be noted however, that if the origin point was translated from the centre of $\boldsymbol{S}_{FG}$ to the lower left hand corner, this object would become detectable without degrading the detection of the other flat-topped objects.

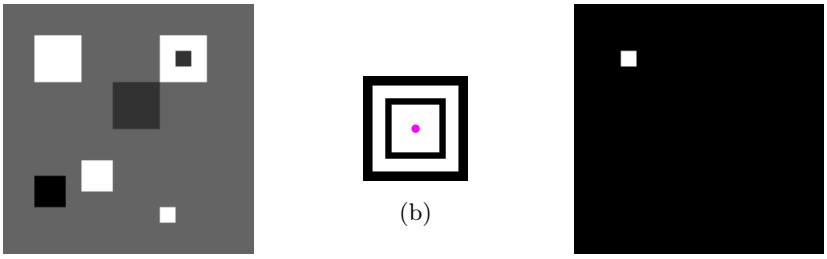
- The bottom left square is similar to the previous two but includes a single, central, pixel with an intensity below that of the background, $t = 50$. Therefore, the dilation of the background, $[\delta_{\boldsymbol{S}_{BG}}](p_3) = 100$, produces a value greater than that of the erosion of the foreground, $[\varepsilon_{\boldsymbol{S}_{FG}}](p_3) = 50$. As a result, it is detectable by neither the UHMT nor the CHMT.

- The bottom right square is almost identical to the top left, with a reduced intensity value of $t = 150$. As this value is greater than that of the background, this object is detected by both the UHMT and CHMT, *i.e.,* $[\text{UHMT}_{\boldsymbol{S}}(\boldsymbol{X})](p_4) = [\text{CHMT}_{\boldsymbol{S}}(\boldsymbol{X})](p_4) = 50$.

An additional, generalised, definition based on Soille's CHMT and UHMT is used as the basis for the HMT in this thesis. In Equation (3.19), the output of the HMT is a binary marker based solely on whether the foreground SE fits and the background SE misses, and therefore detects, an object of interest.

$$[\text{HMT}_{\boldsymbol{S}}(\boldsymbol{X})](x) = \begin{cases} 1, & \text{if } [\delta_{\boldsymbol{S}_{BG}}](x) < [\varepsilon_{\boldsymbol{S}_{FG}}](x) \\ 0, & \text{otherwise} \end{cases} \tag{3.19}$$

An example of using this generalised form of the HMT from Equation (3.19) for performing object detection on a synthetic greyscale image can be seen in Figure 3.10. Similarly to the binary image test shown in Figure 3.7, the SE has been designed to detect 3×3 squares brighter that their background.



(a) (b) (c)

Figure 3.10: Example of a greyscale HMT: **a)** Greyscale test image made up of of shapes with varying intensity and size. **b)** Greyscale composite SE. **c)** HMT of a) using b).

3.2.6 Noise Robustness and the Percentage Occupancy Hit-or-Miss Transform

While it performs well in many applications, noise in an image reduces the reliability of the HMT as an object detection technique. This is due to the underlying assumption that both SEs will fit the object under inspection ideally, meaning just one noisy pixel in the foreground or background can thwart this operation. This is apparent in Figures 3.7, 3.9 and 3.10 where, in several cases, there are white square shapes with a single non-white pixel that are not detected using the standard binary and greyscale HMTs.

There are multiple methods of relaxing the hard constraints of the HMT to detect these imperfect objects including the notion of PO. Here, occupancy refers to the proportion of elements in the support of an SE when translated to a point, x, that lie on the graph or within the sub-graph of an image; *i.e.,* if the occupancy is 100%, the object can be detected using the standard HMT, whereas any occupancy value less than 100% would result in a missed object using a HMT. PO allows for a percentage error in the matching of an SE to an image and led to the development of the POHMT [49]. These PO measures can be implemented by replacing the hard limits of erosions (minimum filters) and dilations (maximum filters) with rank k filters [49]. The output of a rank order filter with rank k within the region of an SE, $\boldsymbol{S}$, translated to a point, x, in an image, $\boldsymbol{I}$, is calculated using the following:

$$[\zeta_{\boldsymbol{S},k}(\boldsymbol{I})](x) = k^{\text{th}} \text{ order} \underset{s\in\boldsymbol{S}}{\text{statistic}}\{\boldsymbol{I}(x+s)\} \tag{3.20}$$

A rank order filter with rank k will return the k^{th} smallest element, or k^{th} order statistic, of an ordered set of data. When $k = 1$, an erosion takes place and when $k = \text{card}(\boldsymbol{S})$ a dilation takes place. Minimum, maximum, and median filters are all examples of rank order filters.

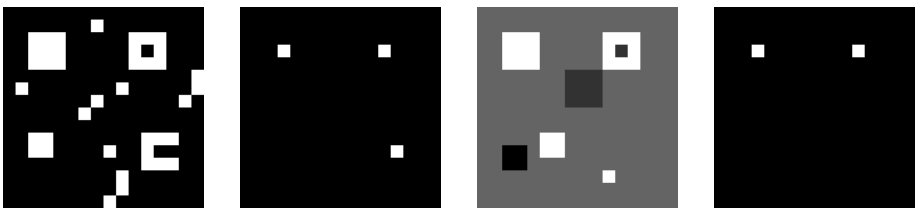
This equivalence between the erosion and dilation and rank order filters greatly simplifies the implementation of morphological operations. The desired rank given some PO value, p, is denoted as k_p and allows for the unbiased specification of rank regardless of the size and shape of an SE, $\boldsymbol{S}$, and the region it describes within an image [49]:

$$k_p = \left\lfloor \left(\frac{p}{100} \text{card}(\boldsymbol{S}) \right) \right\rfloor \tag{3.21}$$

The POHMT can then be defined in terms of the values of two rank order filters and the desired value of k given some value for p, k_p. The first of these operations is equivalent to a relaxed erosion of the foreground, $[\zeta_{\boldsymbol{S}_{FG},k_{100-p}}(I)](x)$, and the second is the relaxed background dilation, $[\zeta_{\boldsymbol{S}_{BG},k_p}(I)](x)$, such that:

$$[\text{POHMT}_{\boldsymbol{S}}(\boldsymbol{X})](x) = \begin{cases} 1, & \text{if } [\zeta_{\boldsymbol{S}_{BG},k_p}(\boldsymbol{I})](x) < [\zeta_{\boldsymbol{S}_{FG},k_{100-p}}(\boldsymbol{I})](x) \\ 0, & \text{otherwise.} \end{cases} \tag{3.22}$$

In Equation (3.22) the output of the POHMT is the set of points in an image, $\boldsymbol{I}$, such that for any $x \in \boldsymbol{I}$ the intensity at rank k_{100-p} of the image values in $\text{supp}(\boldsymbol{S}_{FG})$, when translated to some point, x, is greater than the intensity value at rank k_p of $\text{supp}(\boldsymbol{S}_{BG})$ translated to the same point. An example of the POHMT is shown in Figure 3.11 applied to the images from Figures 3.7 and 3.10 with $p = 70\%$ and $p = 80\%$ respectively. This allows for the detection of objects in the presence of noise as shown in Figure 3.11 where the objects with holes in an otherwise square shape are detected despite the non-ideal shape or intensity. However, this relaxation can also introduce False Positive (FP) detections with improper SE design and choice of rank parameter, k_p, especially in a scene including similar target and non-target objects.



(a) (b) (c) (d)

Figure 3.11: **a)** Binary test image from Figure 3.5a **b)** POHMT of a) using Figure 3.5b and $p = 70\%$. **c**) Greyscale test image from Figure 3.9a **d)** POHMT of c) using Figure 3.9b and $p = 80\%$.

3.2.7 Optimal Parameter Setting using a PO Plot

In order to ascertain the appropriate PO value to achieve favourable performance using the POHMT, query points in an image can be used to create PO plots. The PO plot forms a visual representation of how well the composite SE fits within the image when translated to a given pixel. The occupancy of the foreground and background SEs are calculated and plotted against each other in one of two ways.

The first involves plotting the occupancy of both the foreground and background SEs and finding the point at which they intersect using:

$$PO_{\boldsymbol{S}}(x) = \max_{t \in T}[\min\{PO_{FG_{x,t}}, PO_{BG_{x,t}}\}]. \tag{3.23}$$

Where $PO_{FG_{x,t}}$ is the percentage of pixels within the support of the foreground SE when translated to some pixel, x, which are greater than or equal to each grey level, t, in the set of all grey levels, T. $PO_{BG_{x,t}}$ is the percentage of pixels within the support of the background SE when translated to some pixel, x, which are less than each grey level. The optimal rank order filter to detect a given object at this particular point can be defined using Equation (3.20), where the rank value is calculated using the result of Equation (3.21), *i.e.,* $k_{PO_{\boldsymbol{S}}(x)}$.

The second method for obtaining the optimal PO value involves plotting $PO_{FG_{x,t}}$ against $PO_{BG_{x,t}} \forall t \in T$. The optimal PO value is then calculated as the point where the resultant line intersects the diagonal line given by $PO_{FG} = PO_{BG}$. While both methods inform the optimal PO value, any PO value less than the optimal will result in a positive detection when the composite SE, $\boldsymbol{S}$, is translated to the same pixel, x. The full derivations for occupancy and the resultant PO plots are found in [50].

Figure 3.12 shows an example of using the POHMT with optimal rank order filters, derived from PO plots, for detecting objects in a noisy image. Figure 3.12a shows the clean image and Figure 3.12b shows the same image with white Gaussian noise applied, with $\mu = 0$ and $\sigma^2 = 0.2$ clipped to the range $[0, 1]$. The surface representations of Figures 3.12a and 3.12b are shown in Figures 3.12c and 3.12d respectively. A composite structuring element, $\boldsymbol{S}$, was designed where a 15×15 square was used as the foreground SE and a 17×17 hollow square with a 15×15 hole removed such that $\boldsymbol{S}_{FG} \bigcap \boldsymbol{S}_{BG} = \varnothing$. Translating $\boldsymbol{S}$ to the position denoted by the cross in Figures 3.12a and 3.12b, produces the PO plots shown in Figures 3.12e and 3.12g for the clean test image and Figures 3.12f and 3.12h in the noisy case.

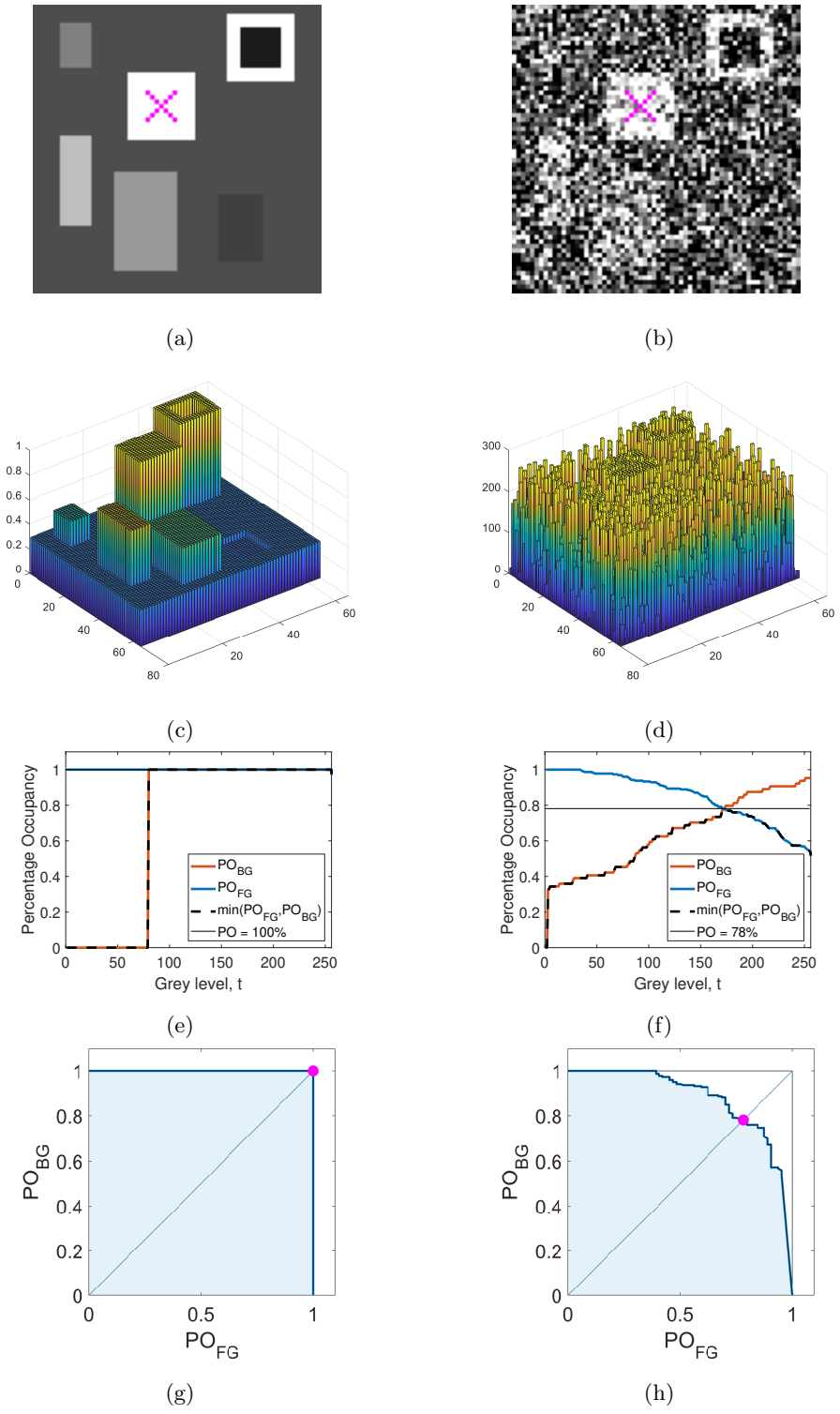


Figure 3.12: **a)** Clean test image with the marked test location. **b)** Noisy test image. **c)** Clean surface. **d)** Noisy surface. **e)** Foreground and background occupancies using an SE centred at the test location in the clean image. **f)** Foreground and background occupancies using an SE centred at the test location in the noisy image. **g)** PO plot in the clean image. **h)** PO plot in the noisy image.

When the designed SE, Figures 3.13a and 3.13b, is centred on the pixel marked with a cross the image where no additive noise is present, as shown in Figure 3.12a, the object can be detected with no relaxation, *i.e.,* a PO value of 100%.

This can be verified by plotting the occupancies of the foreground and background SEs, as in Figure 3.12e; where the orange line represents the occupancy of the background SE, $PO_{BG_{x,t}} \forall t \in T$, and the blue line represents the same for the foreground SE, $PO_{FG_{x,t}} \forall t \in T$. The dashed black line indicates the minimum value of $PO_{FG_{x,t}}$ and $PO_{BG_{x,t}} \forall t \in T$ and is a visual representation of Equation (3.23) where the maximum value of this new line is the optimal PO value and is indicated by the horizontal black line in Figure 3.12e at $PO = 100\%$. Figure 3.12g shows the equivalent PO plot which forms a right angle with a corner at $PO_{FG} = PO_{BG} = 100\%$, selecting any PO_{FG} and PO_{BG} value within the blue shaded area will result in a positive detection, however, this may also cause undesirable FP detections.

Similarly for the noise compromised image in Figure 3.12b, the foreground and background occupancies are plotted in Figure 3.12f with the PO plot shown in Figure 3.12h. In this case, the PO is required to be relaxed to 76% occupancy in order to account for the noise present over the object of interest, as indicated by both the horizontal line in Figure 3.12f and the intersection point in Figure 3.12h represented by the magenta dot.

The occupancy given some SE can be shown as an image, where each pixel value represents the estimated occupancy to register a hit. An example of the occupancy images for the clean and noisy images from Figure 3.12 can be seen in Figure 3.13. In Figure 3.13c, there is a single pixel equal to 1, *i.e.,* has an occupancy of 100%, which corresponds to the centre of the square object marked with a cross in Figure 3.12a. In Figure 3.13d, the value of this same pixel has dropped to 0.78, *i.e.,* 78% occupancy, and would require relaxation to detect using the SE shown in Figures 3.13a and 3.13b, as is the case in Figure 3.13e.

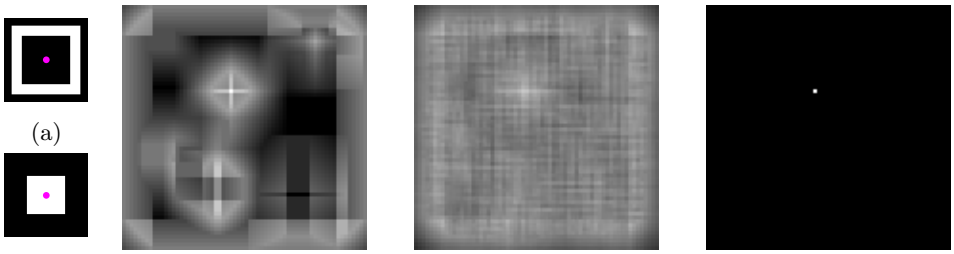


Figure 3.13: POHMT result using the optimal parameters in from Figure 3.12. **a)** $\boldsymbol{S}_{BG}$. **b)** $\boldsymbol{S}_{FG}$. **c)** Occupancy of Figure 3.12a using the composite SE, $\boldsymbol{S}$. **d)** Occupancy of Figure 3.12b using the composite SE, $\boldsymbol{S}$. **e)** POHMT result of Figure 3.12b where $p = 77\%$.

3.3 Multivariate Mathematical Morphology

Any further extension of MM to RGB and higher dimensionality domains in the same vein as the initial binary to greyscale extension is non-trivial. There is no analogue to the surface representation used in greyscale morphology nor is there an intrinsic notion of order in multivariate space. The fundamental issue preventing a simple extension to multichannel images is that MM requires a complete lattice [41] in order for its operators to be accurately defined, a trivial task in scalar variables but a far more ill-posed problem in multivariate mathematics. Defining a complete lattice, *i.e.,* one with an *infimum* and a *supremum*, on multivariate data is a challenging task as there is no generally accepted or unambiguous way to order such data [215]. As a result, no unified morphological framework exists for use in colour or multivariate images. There have, however, been multiple attempts at generalising MM and the HMT in particular to the colour and multivariate image domains. In this section, the fundamentals of colour ordering and how the ordering schemes detailed in [51] can be incorporated into a morphological framework are discussed.

3.3.1 Multichannel Image Formation

Unlike the set mappings in Equations (3.1) and (3.2) for binary and greyscale images respectively, an equivalent set representation for multichannel images does not exist. Consequently, there is no unifying extension of morphological techniques to colour and multivariate images without some additional processing. Instead, each of the pixels, p, in a multivariate, or multi-channel, image, $\boldsymbol{f}$ with m channels can be defined as an m-dimensional vector:

$$\boldsymbol{f}(p) = (f_1(p), f_2(p), ..., f_m(p)) \tag{3.24}$$

or alternatively, multi-channel images can be represented as a stack of m greyscale images.

3.3.2 Multivariate Ordering Relations

As discussed in Section 2.2.2, there are four main categories of multivariate ordering schemes as defined in [51]. Examples of marginal, conditional, and reduced ordering schemes on the same synthetic colour image data are shown here. The fourth ordering scheme, partial ordering, is not considered as it requires the development of some measure of extremeness. In each example, the individual colour channels as well as the composite "pixel" colour they form is displayed before

and after ordering. The pixel values, and by extension the composite colour, are randomly selected using

$$\{(r_i, g_i, b_i) \mid r_i, g_i, b_i \sim \{0.1, 0.2, \ldots, 0.9, 1.0\}\}_{i=1}^{n} \tag{3.25}$$

such that n random colour pixels are created for ordering. These pixel values are sampled uniformly from the relatively small discrete set of values so as to guarantee there will be some repetition in order to highlight how each ordering scheme behaves.

Marginal ordering

Marginal ordering is achieved by ordering each channel independently and an example of an *m-ordering* scheme can be seen in Figure 3.14.

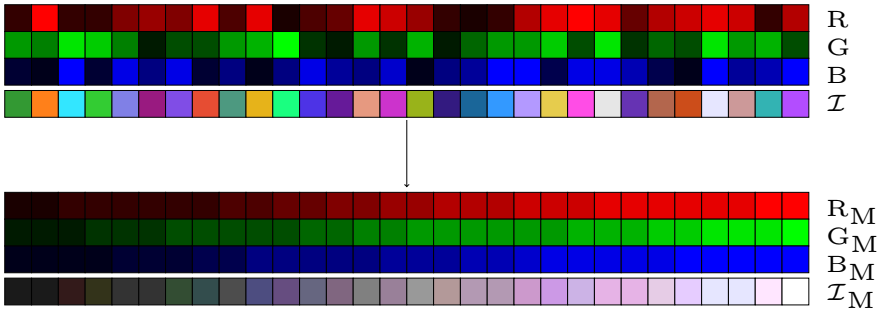


Figure 3.14: Marginal ordering on synthetic colour data.

In Figure 3.14 the top three rows, R, G, and B, represent the red, green, and blue components of each of the input pixels respectively. The fourth row, $\mathcal{I}$, represents the resultant colour of each of these pixels in the RGB colour space.

These channels can be ordered individually using an *m-ordering* scheme, and results in the order shown in $\mathrm{R_M}$, $\mathrm{G_M}$, and $\mathrm{B_M}$, for the red, green and blue channels respectively. Combining each channel into a composite colour pixel results in the colours shown in $\mathcal{I}_\mathrm{M}$.

Ordering the channels individually in this way destroys the inter-channel correlation and introduces false colours. Although an extreme example, as adjacent pixels in an image are likely to have similar colours, if the colour erosion or dilation of this region in an image was taken, the result would be the first (dark grey) or last (white) element in the set of ordered colours respectively, neither of which are present in the original set of input colours.

Conditional ordering

Conditional, or lexicographical, ordering schemes order multivariate pixels using a priority or

condition on their channels which is defined in advance. Any two pixels with equal values in the channel with highest priority, are then ordered based on their values in the subsequent channel(s). For two 3-dimensional points $\boldsymbol{a}$ and $\boldsymbol{b}$, a conditional order can be inferred using the following set of equations,

$$\boldsymbol{a} < \boldsymbol{b} \Leftrightarrow \begin{cases} a_1 < b_1 \text{ or} \\ a_1 = b_1 \text{ and } a_2 < b_2 \text{ or} \\ a_1 = b_1 \text{ and } a_2 = b_2 \text{ and } a_3 < b_3 \end{cases} \tag{3.26}$$

where $\boldsymbol{a} = \{a_1, a_2, a_3\}$ and $\boldsymbol{b} = \{b_1, b_2, b_3\}$. An example of conditional ordering on the same synthetic data used in the marginal scheme from Figure 3.14 can be seen in Figure 3.15, where each colour channel is used in turn in the lexicographical cascade from Equation (3.26) (*i.e.,* prioritising red, then green, then blue).

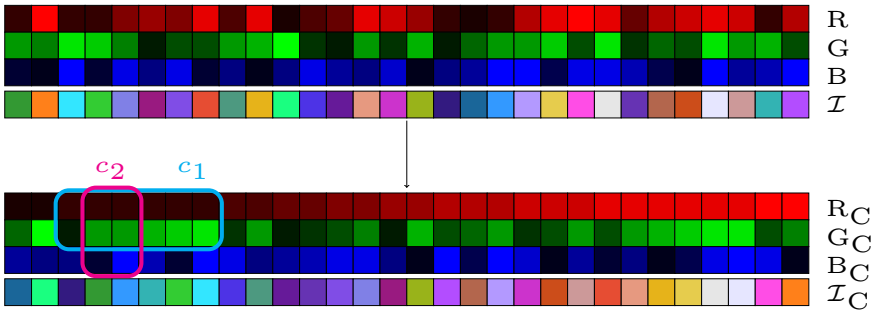


Figure 3.15: Conditional ordering on synthetic colour data.

The top four rows shown in Figure 3.15, R, G, B, and $\mathcal{I}$ represent the unsorted red, green, and blue components and composite RGB colour of each pixel respectively. $\mathrm{R_C}$, $\mathrm{G_C}$, $\mathrm{B_C}$, and $\mathcal{I}_\mathrm{C}$ represent the resultant ordered set following a *c-ordering* approach. The red elements, $\mathrm{R_C}$, are "correctly" ordered from 0.1 to 1.0 as this channel has the highest priority, with the green, $\mathrm{G_C}$, and blue, $\mathrm{B_C}$, components appearing to be more irregular. Exceptions to this irregularity can be seen in areas where the red channel contains repeat elements and alone is insufficient to derive an order. The region indicated by the rectangle c_1 in Figure 3.15 contains six repeat elements in the red channel, as a result, the green values of these pixels are used to derive the order, utilising the second condition in the lexicographical cascade. However, there are also two repeated green values within this region, as seen in the region bounded by c_2, and as a result the blue values are required to determine the order, using the third condition in the lexicographical cascade.

Conditional ordering produces a total order, provided each component in the data is used [216]. However, the resultant order may be arbitrarily decided by only the first few components,

as in Figure 3.15 where the majority of sorting is carried out based on the red channel. Although this method does not introduce false-colours, as is the case in a marginal ordering approach, it does introduce a bias towards the channel which is given highest priority. In the above example, blue-green colours (with low correlation to red) are ordered lowest and would be the result of an erosion; whereas redder colours, orange in the case of Figure 3.15, are ordered highest and would be the result of a dilation. Depending on the context of the image, this may be meaningless and as a result, the priority placed on each channel in the image must be carefully considered.

Reduced ordering

The final example of multivariate ordering considered in this thesis is reduced ordering, which introduces a reference vector, $\boldsymbol{r}$, which seeks to reduce m-dimensional data to a scalar representation which can be easily sorted. For each measurement, $\boldsymbol{x}_i \in \boldsymbol{X}$, some difference metric, $d : \mathbb{R}^m \to \mathbb{R}$, between $\boldsymbol{x}_i$ and $\boldsymbol{r}$ is calculated such that a scalar $d_i = d(\boldsymbol{x}_i, \boldsymbol{r})$ is obtained. Sorting each value of d_i allows an order on $\boldsymbol{X}$ to be inferred where n, m-dimensional, vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n$ are ordered based on $d_1, d_2, ..., d_n$ as follows:

$$\boldsymbol{x}_{\sigma(1)} \leq \boldsymbol{x}_{\sigma(2)} \leq \ldots \leq \boldsymbol{x}_{\sigma(n)} \tag{3.27}$$

where σ is the permutation of the indices $\{1, 2, \ldots n\}$ such that:

$$d_{\sigma(1)} \leq d_{\sigma(2)} \leq \ldots \leq d_{\sigma(n)}. \tag{3.28}$$

An example of reduced ordering of the synthetic colour data is shown in Figure 3.16. In this example a blue pixel, to which every other observation is measured against, is introduced as the reference, $\boldsymbol{r}$, which allows for desirable colours to be prioritised in the output image.

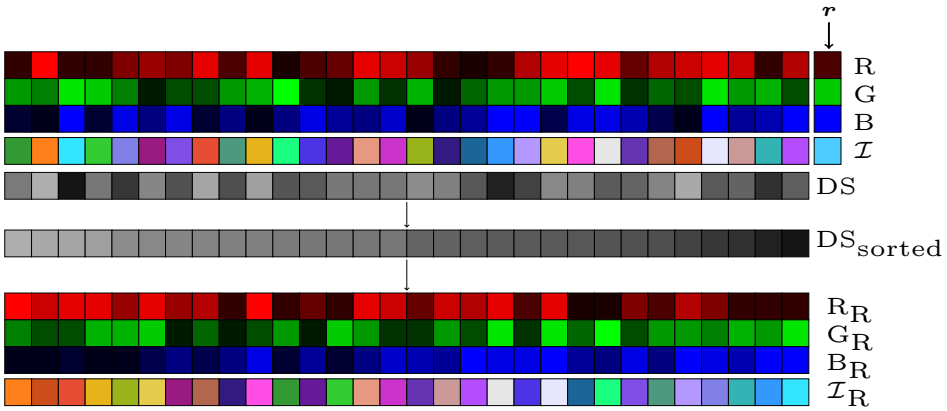


Figure 3.16: Reduced ordering on synthetic colour data.

Two additional rows are introduced alongside the input and ordered data. The first, DS, represents the Euclidean distance, used as the difference metric d, between each of the input colours and the reference and is equivalent to the set of scalar differences d_i. The second is the sorted set of differences which informs the colour order. As seen in Figure 3.16, the original colours are maintained and ordered from orange, which is chromatically opposite to the provided reference, to blue - which was the desired colour.

3.3.3 The Similarity Hit-or-Miss Transform

The SHMT [40, 217] is a colour HMT similar in operation to the method proposed in Chapter 4 of this thesis and is based on the notion of colour similarity to define an ordering. The colour similarity between an image pixel, $\boldsymbol{x} = [x_{_R},\ x_{_G},\ x_{_B}]$, and reference, $\boldsymbol{r} = [r_{_R},\ r_{_G},\ r_{_B}]$, in the RGB space is defined as

$$h_c(\boldsymbol{x}) = \phi_c(\boldsymbol{x}, \boldsymbol{r}) = \mathrm{NI}(\boldsymbol{x}, \boldsymbol{r}) \times \mathrm{chroma}(\boldsymbol{x}, \boldsymbol{r}) \tag{3.29}$$

where the normalised intensity $\mathrm{NI}(\boldsymbol{x}, \boldsymbol{r})$ and chromatic similarity $\mathrm{chroma}(\boldsymbol{x}, \boldsymbol{r})$ between $\boldsymbol{x}$ and $\boldsymbol{r}$ are defined as

$$\mathrm{NI}(\boldsymbol{x}, \boldsymbol{r}) = 1 - \left| \frac{\|\boldsymbol{x}\|}{\|\boldsymbol{x}^{\mathrm{MAX}}\|} - \frac{\|\boldsymbol{r}\|}{\|\boldsymbol{r}^{\mathrm{MAX}}\|} \right| \tag{3.30}$$

and

$$\mathrm{chroma}(\boldsymbol{x}, \boldsymbol{r}) = 1 - \frac{2}{\pi} \cos^{-1} \frac{\left\langle \boldsymbol{x}^{\mathrm{MAX}}, \boldsymbol{r}^{\mathrm{MAX}} \right\rangle}{\|\boldsymbol{x}^{\mathrm{MAX}}\| \|\boldsymbol{r}^{\mathrm{MAX}}\|} \tag{3.31}$$

respectively, where $\boldsymbol{x}^{\mathrm{MAX}}$ and $\boldsymbol{r}^{\mathrm{MAX}}$ are the colour extrema of both the image pixel and the reference colour defined in [217].

With this definition of h-similarity, an h-ordering scheme can be defined as a lexicographical cascade of this h-similarity value, followed by the green, red, and blue values in turn, as shown in Equation (3.32).

$$\boldsymbol{x}_i < \boldsymbol{x}_j \Leftrightarrow \begin{cases} \phi_c(\boldsymbol{x}_i, \boldsymbol{r}_i) < \phi_c(\boldsymbol{x}_j, \boldsymbol{r}_j) \text{ or} \\ \phi_c(\boldsymbol{x}_i, \boldsymbol{r}_i) = \phi_c(\boldsymbol{x}_j, \boldsymbol{r}_j) \\ \textit{and} \\ \begin{cases} \boldsymbol{x}_i^G < \boldsymbol{x}_j^G \text{ or} \\ \boldsymbol{x}_i^G = \boldsymbol{x}_j^G \text{ and } \boldsymbol{x}_i^R < \boldsymbol{x}_j^R \text{ or} \\ \boldsymbol{x}_i^G = \boldsymbol{x}_j^G \text{ and } \boldsymbol{x}_i^R = \boldsymbol{x}_j^R \text{ and } \boldsymbol{x}_i^B < \boldsymbol{x}_j^B \end{cases} \end{cases} \tag{3.32}$$

An example of the h-ordering scheme is shown in Figure 3.17 using the same set of image and reference colours from Figures 3.14 to 3.16.

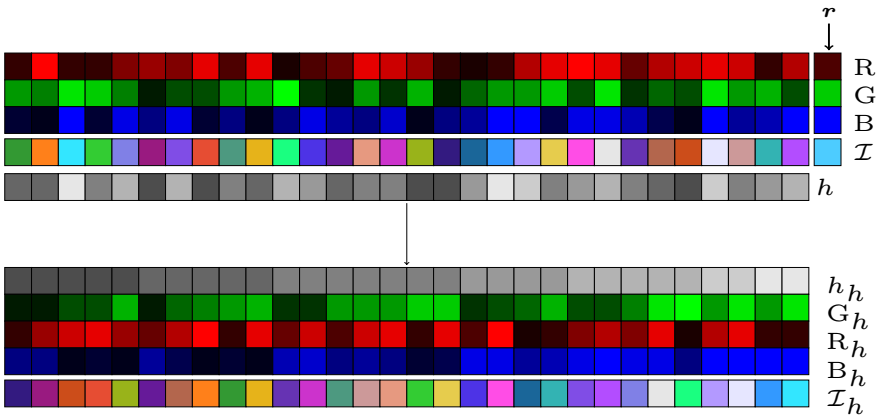


Figure 3.17: h-similarity ordering on synthetic colour data.

For an image, $\boldsymbol{I}$, and an SE, $\boldsymbol{S}$, the SHMT is defined using h-similarity as

$$\text{SHMT}(\boldsymbol{I};\boldsymbol{S}) = \left\{ h^s_{\text{similarity}} \left(\varepsilon_{h^s_{\text{similarity}}} \boldsymbol{I}(x) \right) = 1, x \in \boldsymbol{E} \right\} \tag{3.33}$$

for all points x in $\boldsymbol{I}$, where $\mathbf{E}$ is the 2D support of the image in $\mathbb{Z}^2$ and $h^s_{\text{similarity}}$, and $\varepsilon_{h^s_{\text{similarity}}}$ are the colour similarity and colour similarity-based erosion respectively. Using this definition, the result of SHMT is the set of pixels where the resultant h-similarity component $\varepsilon_{h^s_{\text{similarity}}}$, or the minimum similarity between $\boldsymbol{I}$ and $\boldsymbol{S}$ within the support of $\boldsymbol{S}$ $\forall s \in \boldsymbol{S}$, is equal to 1. A variant of the SHMT, the SHMT_η, is also described and is defined as

$$\text{SHMT}_\eta(\boldsymbol{I};\boldsymbol{S}) = \left\{ h^s_{\text{similarity}} \left(\varepsilon_{h^s_{\text{similarity}}} \boldsymbol{I}(x) \right) \geq \eta, x \in \boldsymbol{S} \right\} \tag{3.34}$$

where the result of taking the SHMT_η of $\boldsymbol{I}$ using $\boldsymbol{S}$ is the set of pixels where the $h^s_{\text{similarity}}$ component of the erosion, $\varepsilon_{h^s_{\text{similarity}}}$, is greater than or equal to η.

3.4 Hyperspectral Imaging

Hyperspectral imaging is a technique for collecting information about an imaged scene from across the electromagnetic spectrum, Figure 3.18, and can be considered as a combination of digital imaging with spectroscopy.

RGB images captured by conventional digital cameras are formed of three channels representing the red, green and blue components. These three components are combined to

form a colour picture, mimicking the human visual system's interpretation of light as colour. Each channel is generated by the sensor response to incident light from a wideband region of the electromagnetic spectrum defined by the spectral sensitivity of the colour filters used.

Unlike more traditional, greyscale and RGB sensors, hyperspectral and multispectral sensors capture information from across the electromagnetic spectrum, Figure 3.18, commonly in the Visible and Near-InfraRed (VNIR) range. Multispectral images often contain tens of channels with hyperspectral images often being made up of hundreds or even thousands of individual channels, each captured at a different, narrowband, region of light, as displayed in Figure 3.19, often defined by some centre wavelength, λ_c. The technology underpinning hyperspectral imaging was developed as a portable field spectrometer to support the LandSat programme in identifying mineral deposits that showed little colour variation in the four band Multispectral Scanner instrument [218]. The benefits of using HSI, and to a lesser extent MSI, are twofold; firstly, the colour resolution is greatly increased compared to RGB imagery, allowing for disambiguation of materials with similar colours but deviations in regions of their spectrum; secondly, such sensors can be sensitive to wavelengths of light beyond the range of human vision, allowing for additional information to be extracted from an imaged scene. In the remainder of this thesis, the notion of dimensionality refers to the number of channels, N, in a given image, being a vector of N-dimensionality when considered as a point in space.

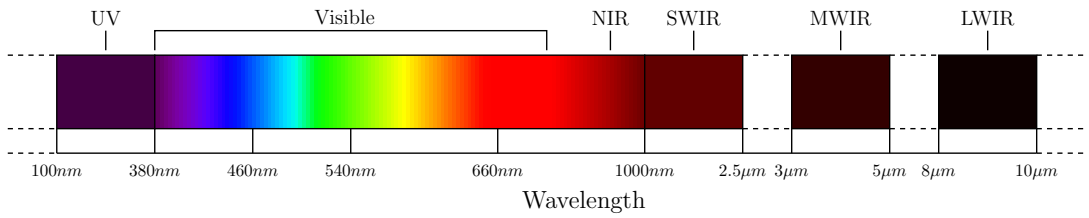


Figure 3.18: Common Coverage of the Electromagnetic Spectrum by Electro-Optical Imaging

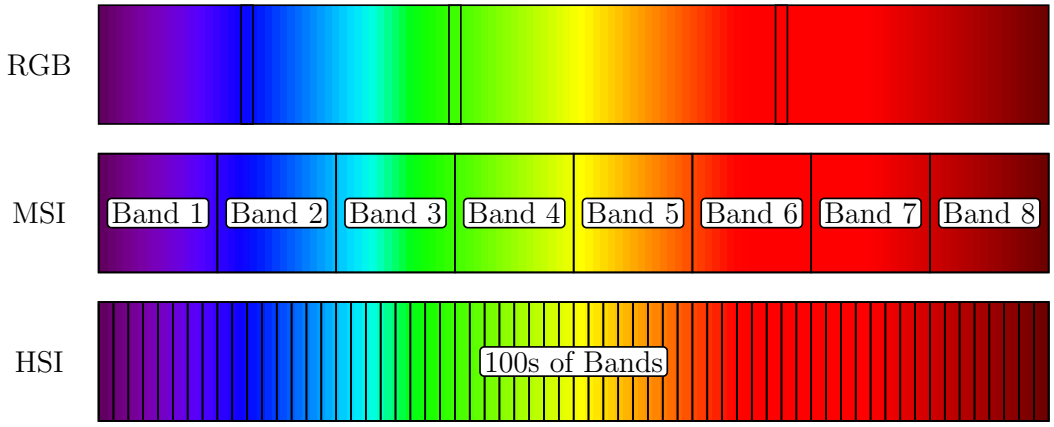


Figure 3.19: Example Electromagnetic Spectrum Sensitivities of Image Modalities

3.4.1 Image Capture

There are five common methods of capturing hyperspectral images, each defined by their use of spatial and spectral sampling and multiplexing techniques and how they address the trade-off between higher resolution and exposure time. Each method is represented in Figure 3.20 and the advantages and disadvantages of each technique are described in the following sections.

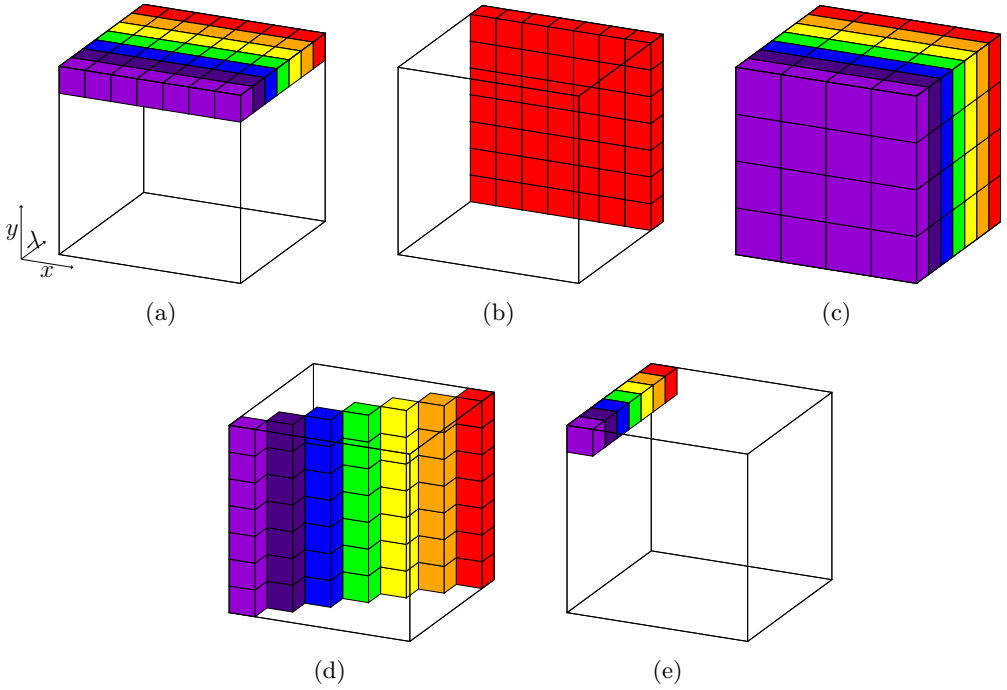


Figure 3.20: Common Methods of Hyperspectral Image Capture: **a)** Pushbroom Scanning. **b)** Spectral Scanning. **c)** Snapshot (non-scanning). **d)** Spatio-Spectral Scanning. **e)** Whiskbroom Scanning.

Pushbroom

Pushbroom (Figure 3.20a) sensors use an optical grating which isolates a single horizontal line of incoming light, this is then diffracted and focussed using prisms and other calibrated optical elements onto a focal plane array. This method captures a single row of spatial pixels and the complete spectrum of each in a single frame. By translating either the camera or the scene and capturing each successive row of pixels, the full image is obtained.

Pushbroom sensors are common in hyperspectral acquisition as they can use the natural motion present in many conveyor-based industrial systems. Similarly, aerial and satellite-based hyperspectral imaging platforms often use pushbroom sensors and exploit the inherent motion of the platform in relation to the scene. Since all wavelengths are captured simultaneously for each row of pixels, the variable sensitivities of sensor materials, such as silicon (Si) or indium gallium arsenide (InGaAs), to different wavelengths of light cannot be accounted for and as such often requires correction via post-processing. Additionally, the exposure time must be linked with the speed of the translation of the sensor, or scene, in order to create square pixels.

Spectral scanning

Spectral scanning (Figure 3.20b), or area scanning, sensors capture the full spatial scene at a single wavelength in each frame. This can be achieved using a sensor which is sensitive to the whole spectrum combined with a bank of filters, each tuned to one of the set of desired wavelengths. By exposing the image through each filter, these single wavelength frames can then be stacked to form the full image.

Such sensors provide a simple and intuitive approach, as each wavelength image is captured in a single exposure. As such, the individual exposures can be tuned separately and both the camera and scene can remain stationary. However, as a result, spectral scanning systems can be limited by the number of wavelengths they can capture. Switching between discrete filters to obtain higher spectral resolutions also increases the total capture time.

Snapshot

Snapshot (Figure 3.20c) spectral cameras capture the entirety of the spatial and spectral information present in a scene in a single exposure. This is achieved by allocating portions of the sensor to capture both spatial and spectral information concurrently which is commonly achieved via the use of DMDs [219]. This sub-division of the focal plane array leads to decreased spatial resolution proportional to any increase in spectral resolution. Alternatively, CSI techniques such as CASSI capture a compressed snapshot but requires later reconstruction [185].

Spatio-spectral scanning

Spatio-spectral scanning (Figure 3.20d) is similar to pushbroom scanning, however, rather than capturing the full spectrum of a single row of pixels, the spectral information captured is of multiple adjacent rows of pixels. This can be achieved by dispersing the light using a slit spectroscope [220], or through filtering techniques such as linear variable filters [221]. Since spectral information is distributed across the spatial scene, further processing is required to translate each set of exposures into a hyperspectral image. Depending on the quality of the filter and the resolution the focal plane array, the resultant spatial and spectral resolutions can vary.

Whiskbroom

Point scanning, or whiskbroom, techniques (Figure 3.20e) capture the whole spectrum of a single point before passing over a scene in a raster scan and repeating this operation at each point in order to form a full hyperspectral image. This operation is similar to a traditional spectrometer being passed over a two dimensional plane allowing for larger spatial and spectral resolutions to be realised at the cost of an increased total capture time. The necessity to move in two dimensions requires much more precise calibration and localisation of pixels in order to represent the scene accurately. The AVIRIS sensor is an example of a hyperspectral imager which operates in a whiskbroom mode [222].

3.5 Dimensionality Reduction

3.5.1 Dimensionality Reduction Algorithms

Dimensionality Reduction techniques can be readily applied to hyperspectral images given their high dimensionality and tendency to exhibit significant inter-band correlation. Most, if not all, algorithms seek to reduce the dimensionality of input data by representing it in a more informative, uncorrelated, subspace which requires fewer coefficients, K, than observations, L, to adequately describe the underlying signal structure and variability. Alternatively as a hyperspectral image can be represented by a finite set of Endmembers, these Endmembers can be estimated and extracted, similar to the Blind Source Separation or *cocktail party* problem. DR in this form seeks to represent the data as a set of spectra and corresponding abundances - along with residual noise or estimation error.

As discussed in Section 2.4, DR has been used in conjunction with various TD algorithms in order to increase their accuracy by simplifying and augmenting the search space for targets

of interest. In this section, the four techniques for spectral dimensionality reduction that are used in the contributions of this thesis are presented. In the following sections, an $M \times N \times L$ hyperspectral image, $\mathbf{X}$, is represented in a reshaped matrix form, $\boldsymbol{X}$, of size $MN \times L$, such that each $1 \times L$ pixel of $\mathbf{X}$ forms a single row, $\boldsymbol{x}$, in $\boldsymbol{X}$.

3.5.1.1 Principal Component Analysis

PCA [75] is a classical method of DR. It seeks to remap highly correlated data into an uncorrelated space using a set of optimal orthogonal basis vectors calculated from the input data where the remapped, or transformed, data is then represented as a set of PCs. These basis vectors are the eigenvectors of the input data, ordered by the magnitude of the corresponding eigenvalue. Each eigenvalue represents the variance of the original data characterised by its eigenvector, akin to the proportion of the data that can be explained by a particular component. By retaining the first K components, where $K < L$, the data can be compressed. By performing PCA and remapping into an uncorrelated space, potential targets can become more prominent when compared to their background.

There are multiple methods for performing PCA through both iterative and non-iterative algorithms, two are discussed here; EVD and NIPALS.

A) Eigenvalue Decomposition

The most common method for performing PCA is to use an EVD. First, the covariance matrix, $\boldsymbol{\Sigma}$, of an input image, $\boldsymbol{X}$, is calculated as follows

$$\begin{aligned} \bar{\boldsymbol{x}} &= \frac{1}{MN} \sum_{i=1}^{MN} \boldsymbol{x}_i, \\ \boldsymbol{y}_i &= \boldsymbol{x}_i - \bar{\boldsymbol{x}}, \\ \boldsymbol{\Sigma} &= \frac{1}{MN-1} \boldsymbol{Y}^T \boldsymbol{Y}, \end{aligned} \tag{3.35}$$

where $\boldsymbol{Y}$ is a matrix containing the mean adjusted spectra of $\boldsymbol{X}$, $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ are the i^{th} rows of $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively, and x_{i_l} is the l^{th} element of $\boldsymbol{x}_i$ corresponding to the intensity of that pixel at the l^{th} wavelength of light. The eigenvalues, $\boldsymbol{\lambda}$, and eigenvectors, $\boldsymbol{u}$, of the square covariance matrix, $\boldsymbol{\Sigma}$, can be calculated using

$$\boldsymbol{\Sigma}\boldsymbol{u} = \lambda \boldsymbol{u}, \tag{3.36}$$

$$\det(\boldsymbol{\Sigma} - \boldsymbol{\lambda}\boldsymbol{I})\boldsymbol{u} = 0. \tag{3.37}$$

Equation (3.35) can be re-written in matrix form as $\boldsymbol{\Sigma U} = \boldsymbol{U\Lambda}$ or alternatively as

$$\boldsymbol{\Sigma} = \boldsymbol{U\Lambda U}^T, \tag{3.38}$$

where the matrix $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{\Sigma}$, *i.e.,* $\boldsymbol{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, ..., \lambda_L\}$ and the matrix $\boldsymbol{U}$ contains the related eigenvectors $[\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_L]$.

The eigenvalues in $\boldsymbol{\Lambda}$ are ordered such that $\lambda_1 > \lambda_2 > ... > \lambda_L$, hence the first K largest eigenvalues correspond to the first K eigenvectors which represent almost all the significant information contained in the data. These eigenvectors can be retained for DR purposes *i.e.,*

$$\boldsymbol{V}_{PCA} = \boldsymbol{U}_K \tag{3.39}$$

where, $\boldsymbol{U}_K$ contains $[\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_K]$

$$\hat{\boldsymbol{X}}_{PCA} = \boldsymbol{V}_{PCA}^{-1}\boldsymbol{X}_{2D} \tag{3.40}$$

Using Equation (3.40) images and the spectral information of targets that may be present in the scene can be projected into the same reduced dimensionality subspace for processing.

B) Non-Linear Iterative Partial Least Squares

In some cases, such as those where the desired number of components to retain is known, it is unnecessary and therefore preferable to avoid calculating every PC as in an EVD. In these cases iterative techniques can be used to calculate each successive PC in turn, until the required number, K, has been reached. One method to achieve this is by using the NIPALS algorithm which consists of the decomposition $\boldsymbol{X} = \boldsymbol{TP}^T$ where $\boldsymbol{X}$ is some mean-centred matrix and the columns of $\boldsymbol{T}$ are the scores and the columns of $\boldsymbol{P}$ are the loadings, $\boldsymbol{T}$ is an $MN \times L$ matrix and $\boldsymbol{P}$ is $L \times L$. These scores and loadings are found iteratively; scores represent the original data projected into the PC space, loadings are the contribution of each of the original variables to this PC and are mathematically similar to eigenvectors. The iterator, i, is initialised as $i = 1$ and $\boldsymbol{X}_i = \boldsymbol{X}$, an initial guess of the i^{th} loading $\boldsymbol{t}_i$ is made as any column of $\boldsymbol{X}_i$ before an initial set of loadings, $\boldsymbol{p}_i$, can be calculated using

$$\boldsymbol{p}_i = \frac{\boldsymbol{X}_i^T\boldsymbol{t}_i}{\boldsymbol{t}_i^T\boldsymbol{t}_i}. \tag{3.41}$$

The loadings are then normalised using

$$\hat{\boldsymbol{p}}_i = \frac{\boldsymbol{p}_i}{\|\boldsymbol{p}_i^T \boldsymbol{p}_i\|}, \tag{3.42}$$

it can be noted that this step produces a vector mathematically identical to the equivalent eigenvector using an EVD. The scores are then calculated as

$$\boldsymbol{t}_i = \frac{\boldsymbol{X}_i \hat{\boldsymbol{p}}_i}{\hat{\boldsymbol{p}}_i^T \hat{\boldsymbol{p}}_i}. \tag{3.43}$$

While $i \leq K$, Equations (3.41) to (3.43) are repeated until a convergence is reached for each value of $\hat{\boldsymbol{p}}_i$ and $\boldsymbol{t}_i$. After this convergence point is reached, the i^{th} eigenvalue can be calculated as $\lambda_i = \boldsymbol{t}_i^T \boldsymbol{t}_i$ and the corresponding eigenvector is simply $\hat{\boldsymbol{p}}_i$. In each successive iteration, $\boldsymbol{X}_{i+1} = \boldsymbol{X}_i - \boldsymbol{t}_i \boldsymbol{p}_i^T$ and Equations (3.41) to (3.43) are again repeated to convergence, continuing until $i = K$. When $i = K$ the loadings, $\boldsymbol{P}$, and scores, $\boldsymbol{T}$, are calculated as

$$\begin{aligned} \boldsymbol{T} &= \begin{bmatrix} \boldsymbol{t}_1 & \boldsymbol{t}_2 & \cdots & \boldsymbol{t}_K \end{bmatrix}, \\ \boldsymbol{P} &= \begin{bmatrix} \hat{\boldsymbol{p}}_1 & \hat{\boldsymbol{p}}_2 & \cdots & \hat{\boldsymbol{p}}_K \end{bmatrix}. \end{aligned} \tag{3.44}$$

The loadings matrix $\boldsymbol{P}$ can be used in a similar manner to the matrix of eigenvectors from an EVD as in Equations (3.39) and (3.40) by projecting a measurement, $\boldsymbol{X}$ into the reduced PCA space

$$\boldsymbol{V}_{NIPALS} = \boldsymbol{P} \tag{3.45}$$

$$\hat{\boldsymbol{X}}_{NIPALS} = \boldsymbol{V}_{NIPALS}^{-1} \boldsymbol{X} \tag{3.46}$$

In testing, both the EVD and NIPALS algorithms produced PCs with identical magnitudes but some which exhibited opposite polarity, as orthogonality can take one of two directions. The EVD has no need to converge and is therefore faster while producing minimal error. For these reasons, in the remainder of this thesis, only the EVD was used to perform PCA-based DR. A simple example of the DR offered using PCA is shown in Figure 3.21 where a pseudo-colour hyperspectral image is mapped into the PC space.

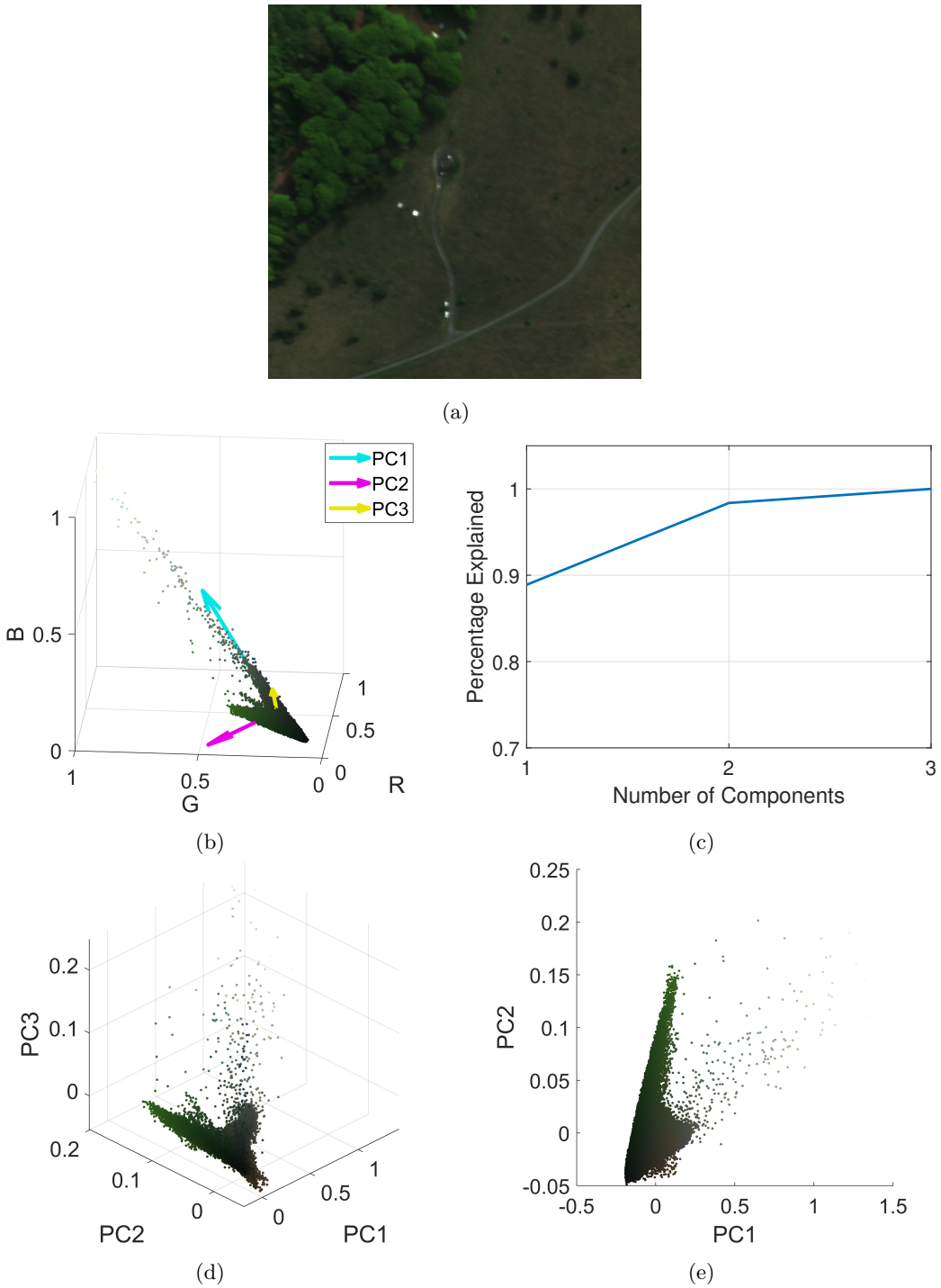
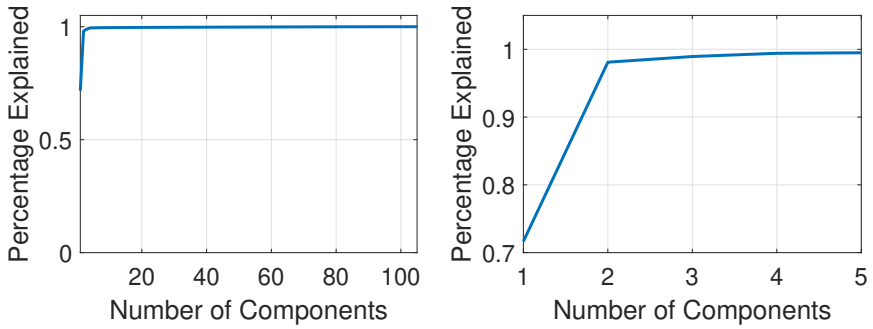


Figure 3.21: Example of using PCA to perform DR on an RGB image. **a)** Pseudo-RGB representation of the "OP7 1" test image Figure A.3.1a. **b)** 3D scatter plot of colour values with overlaid PCs (Eigenvectors). **c)** Percentage data variance explained by each PC. **d)** Colour data remapped to new PC axes. **e)** 2D scatter plot of reduced dimensionality data.

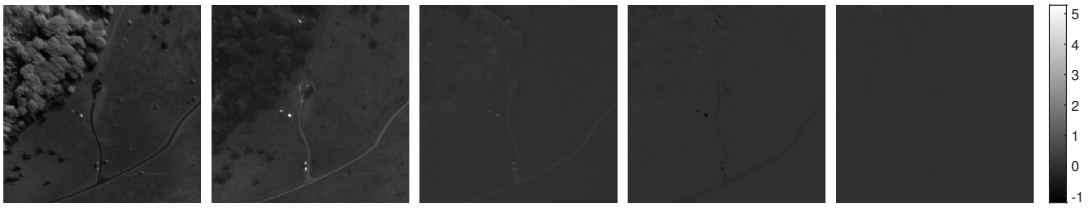
The image in Figure 3.21a can be represented as a 3-Dimensional scatter plot, where each

point is representative of the proportion of red, green, and blue that make up each pixel (Figure 3.21b). Performing PCA on this image produces a set of 3 PCs which, as discussed above, represent the eigenvectors of the data, along which the distribution of the data is optimally represented. These PCs are overlaid onto the scatter plot in Figure 3.21b. The percentage of the original data explained by each PC is shown in Figure 3.21c, with the first PC represents 89% of the variance, and the first two PCs representing 98% of the variance. The original data can be remapped to these PCs, where each axis now represents the value in the PC space as opposed to the original RGB image domain, as shown in Figure 3.21d. As the third principal component represents $< 2\%$ of the original variance exhibited in the data, it can be discarded resulting in a dataset 66.67% the size with negligible loss in quality, going from three dimensions (Figure 3.21d) to two (Figure 3.21e).

Using a hyperspectral image of the same scene as Figure 3.21, the real benefit of DR can be observed where, as visible in Figure 3.22a, only four components are required to represent approximately 99% of the variance in the image, a substantial reduction compared to the original 105 bands at 4% the size of the original dataset. Figure 3.22b shows the first five PC images produced by performing PCA on the OP7 hyperspectral image (Appendix A.3.1) representing 99.48% of the image variance.



(a)



(b)

Figure 3.22: The first 5 Principal Components of the "OP7 1" test image (Figure A.3.1a) calculated using the EVD.

PCA offers a method of calculating a set of optimised basis vectors which concentrate

large amounts of variance in the first few PCs, allowing for high compression rates. The PCs can project data into a subspace exploiting the variance within the data meaning targets of interest, which are often dissimilar to their backgrounds, are highlighted and can become more visible within this subspace. These PCs are optimal as they are calculated on a per-image basis. Therefore, unlike more conventional compression techniques such as the DCT where such compression basis vectors are known, the PCs are only valid for the image from which they were created. This means that if multiple images are processed, PCs must be calculated for each of them.

3.5.1.2 Maximum Noise Fraction

The MNF transform [79] is similar in operation to PCA but additionally accounts for the noise present in input data [81]. Rather than ordering the PCs of an input image, $\boldsymbol{X}$, by their variance, as in PCA, they are instead sorted by their estimated SNR. In MNF, it can be assumed that the covariance of the data, $\boldsymbol{\Sigma}$, is a sum of the covariance of the signal, $\boldsymbol{\Sigma}_{\mathrm{s}}$, and the covariance of noise, $\boldsymbol{\Sigma}_{\mathrm{n}}$, *i.e.,* $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\mathrm{s}} + \boldsymbol{\Sigma}_{\mathrm{n}}$. The MNF transform seeks to maximise the calculated eigenvalues with respect to the estimated SNR.

First, the noise present in $\boldsymbol{X}$ is estimated by taking the difference between adjacent pixels, *i.e.,* $\boldsymbol{X}_n = [\boldsymbol{x}_1 - \boldsymbol{x}_2 \quad ... \quad \boldsymbol{x}_{MN-1} - \boldsymbol{x}_{MN}]$ [79]. In areas of uniform spectral content any discrepancies are usually attributable to noise, $\boldsymbol{X}_n$ is then used to calculate the noise covariance matrix, $\boldsymbol{\Sigma}_n$.

In order to account for the estimated noise, the original data is noise-whitened by first taking the singular value decomposition, $\boldsymbol{\Sigma}_{\mathrm{n}} = \boldsymbol{U}_n \boldsymbol{S}_n \boldsymbol{V}_n^T$ where $\boldsymbol{U}_n$ can be viewed as the set of whitening basis vectors and the noise whitened data matrix can be expressed as $\boldsymbol{X}_W = \boldsymbol{U}_n^T \boldsymbol{X}$. The covariance matrix of the noise-whitened data is equivalent to $\boldsymbol{\Sigma}_{\mathrm{s}}$ and thus by replacing $\boldsymbol{\Sigma}$ from Equation (3.38) with $\boldsymbol{\Sigma}_{\mathrm{s}}$, the eigenvalues and corresponding eigenvectors of $\boldsymbol{\Sigma}_{\mathrm{s}}$ can be found and an input image can be projected into the MNF space in a similar fashion to PCA using

$$\hat{\boldsymbol{X}}_{MNF} = \boldsymbol{V}_{MNF}^{-1} \boldsymbol{X}_{2D}, \tag{3.47}$$

where $\boldsymbol{V}_{MNF}$ consists of the first K eigenvectors calculated using Equations (3.38) and (3.39). An example of MNF-based DR can be seen in Figure 3.23 where the first five MNF components are shown.

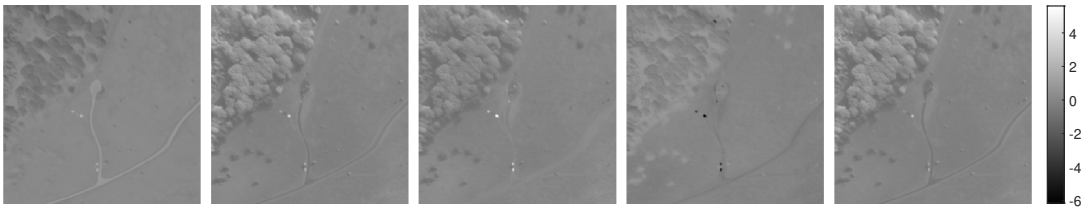


Figure 3.23: The first 5 Principal Components of the "OP7 1" test image (Figure A.3.1a) calculated using the MNF method.

By accounting for the noise in this way, its impact on the resultant PCs is lessened, however cannot be fully removed as $\mathbf{\Sigma}_{\mathrm{n}}$ can only be estimated in most cases. This process of reducing the noise prior to calculating the PCs can be viewed as two separate PCA operations computed in turn, which will have an additional impact on the memory required and computation time.

3.5.1.3 Folded Principal Component Analysis

With both PCA and MNF, as well as many other PCA-like methods, it is necessary to compute the full covariance matrix $\mathbf{\Sigma}$. This covariance matrix is of size $L \times L$ where L is equal to the number of spectral bands in an image. Therefore, for images with high spectral resolution it can be computationally expensive and time-consuming to compute. In order to circumvent this challenge, FPCA [85] seeks to reduce the size of the covariance matrix and also incorporate the correlation within spectra into the calculation. In order to perform FPCA, each of the MN mean-centred spectral vectors, $\bar{\boldsymbol{x}}$, are folded into an $H \times W$ matrix, $\boldsymbol{A}$, where $H \times W = L$ and $H, W \in \mathbb{N}_1$. A partial covariance matrix can be calculated as $\mathbf{\Sigma} = \boldsymbol{A}^T \boldsymbol{A}$ and using each of these MN partial covariance matrices the full covariance matrix, $\mathbf{\Sigma}_{\mathrm{FPCA}}$, can be calculated as

$$\mathbf{\Sigma}_{\mathrm{FPCA}} = \frac{1}{MN} \sum_{i=1}^{MN} \boldsymbol{A}_i^T \boldsymbol{A}_i \tag{3.48}$$

Additional spectra can then be folded using the same H and W and projected using the eigenvectors produced from $\mathbf{\Sigma}_{\mathrm{FPCA}}$, before being unfolded again to be processed in the FPCA domain. An example of the first five PCs calculated using FPCA are shown in Figure 3.24

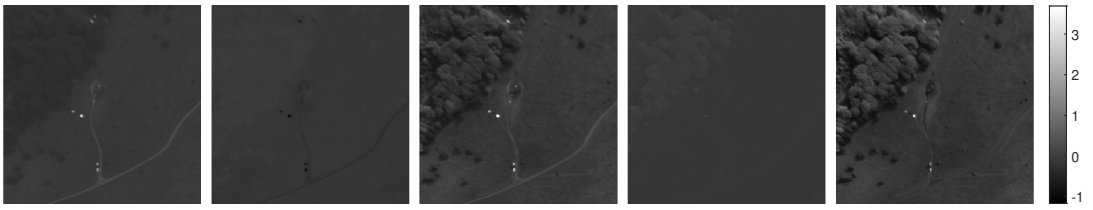


Figure 3.24: The first 5 Principal Components of the "OP7 1" test image (Figure A.3.1a) calculated using the Folded-PCA algorithm.

As with PCA, FPCA projects data into a subspace based on variance and can be useful for increasing the visibility of targets in a given image. As only a partial covariance matrix is calculated the memory requirements for this algorithm are low when compared to PCA and MNF. In testing however, the addition of folding and unfolding operations increased the total computation time when compared to both PCA and MNF. As well as this, the definition of the folding parameters, H and W, were closely linked to the efficacy of the FPCA algorithm.

3.5.1.4 Independent Component Analysis

ICA is a common method for performing Blind Source Separation used in DR. Unlike PCA, MNF or FPCA, ICA seeks to separate an ensemble of mixed signals into a set of finite distinct sources, or ICs, and is achieved by maximising their statistical independence [92]. Hyperspectral images are composed of a weighted sum of a set of finite pure spectra, or Endmembers, and as a result can be decomposed using ICA. There are multiple algorithms used to calculate the ICs of a set of data, two of the most well used are the FastICA [223] algorithm and the Joint Approximation Diagonalization of Eigen-matrices (JADE) algorithm [224]. In the results shown in this thesis, the FastICA algorithm is used instead of the JADE algorithm as it reached convergence both faster and more reliably.

Three approaches for using ICA within the scope of DR are presented in [94], the first of which, *ICA-DR1*, is used as the implementation of ICA-based DR in this thesis. In order to perform ICA-based DR, the number of ICs required to represent the data needs to be estimated. This is achieved by using the notion of VD [225] which estimates the number of spectrally distinct sources in the image, specifically using the Harsanyi–Farrand–Chang (HFC) method [144]. First, an estimate of p sources is obtained using the HFC method, FastICA is then used to obtain $2p$ ICs, or K where $2p < K$ and K is the desired number of dimensions. Doubling the value of estimated independent sources is found to be a good empirical upper bound to account for small and rare targets within the scene [94, 95]. The calculated ICs are not always returned in order of importance, arguably a pre-requisite for any DR scheme, and

as a result a method of determining the importance of each IC is required. In the *ICA-DR1* approach, this is achieved by ordering each of the ICs based on an estimation of their negentropy, with the approximate negentropy of the i^{th} IC, $J(\text{IC}_i)$ calculated as

$$J\left(\text{IC}_i\right) = \left(\frac{1}{12}\right)\left[\kappa_i^3\right]^2 + \left(\frac{1}{48}\right)\left[\kappa_i^4 - 3\right]^2, \tag{3.49}$$

where κ_i^3 and κ_i^4 are the sample means of the third (skewness) and fourth (kurtosis) order statistics of each IC respectively and are calculated as

$$\kappa_i^3 = \frac{1}{MN}\Sigma_{j=1}^{MN}(z_{i_j})^3, \tag{3.50}$$

$$\kappa_i^4 = \frac{1}{MN}\Sigma_{j=1}^{MN}(z_{i_j})^4. \tag{3.51}$$

In Equations (3.50) and (3.51) z_{i_j} is the j^{th} pixel of the i^{th} IC, $\boldsymbol{z_i}$ [95]. Using Equation (3.49) it is possible to order the calculated ICs with the first K components, corresponding to the K largest negentropy values, being used for DR, forming a transform matrix, $\boldsymbol{V}_{ICA}^{-1}$.

$$\hat{\boldsymbol{X}}_{ICA} = \boldsymbol{V}_{ICA}^{-1}\boldsymbol{X}_{2D} \tag{3.52}$$

An example of ICA-DR with the first five ICs can be seen in Figure 3.25.

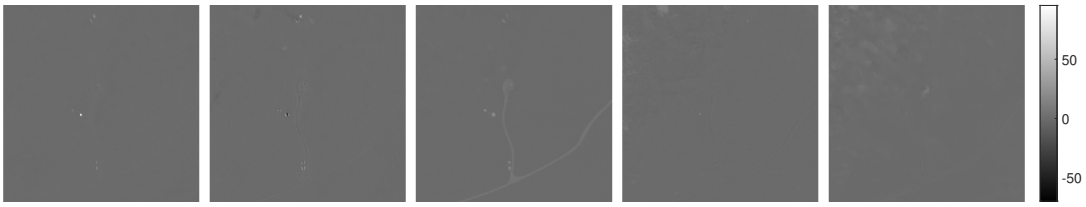


Figure 3.25: The first 5 Independent Components of the "OP7 1" test image (Figure A.3.1a) calculated using the FastICA algorithm ordered using negentropy.

PCA and MNF are both classified as second order statistics-based transforms which can be insufficient in some applications [95]. ICA preserves higher order statistics. While it is possible that second-order statistics may be insufficient in preserving such characterising information, this has not been the case in the results presented in this thesis, as discussed in detail in Chapter 5. Although it performs favourably when compared to other ICA algorithms such as JADE, FastICA is much slower than the other methods for DR listed here. This is another important consideration in its choice in any practical application.

3.5.1.5 Band Selection

Hyperspectral band selection is similar to DR in that it seeks to reduce the number of dimensions, or channels, required to represent the image using statistical inference. However, the manner in which each technique achieves this greatly differs. In DR, the input image is remapped to a more informative basis using a lossy transform, essentially destroying the original structure provided in the set of adjacent channels [103]. Additionally, as these features are created on an *ad hoc* basis for every image individually, there is limited ability to interpret which of the original features were most informative. In contrast, band selection techniques select the most informative set of the original channels based on some measure and use these to reduce the number of channels required to represent the original data.

One common method for performing band selection is the Ward's Linkage strategy using Mutual Information (WaLuMI) algorithm [226, 227]. WaLuMI groups the spectral bands of a hyperspectral image based on the MI exhibited in each. The number of groups, K, is pre-specified, and a representative example from each group is selected under the assumption that each will exhibit a large degree of statistical independence and the final K bands from each of the K groups will be representative of the whole scene. A similar method using the KL divergence in place of MI, Ward's Linkage strategy using Divergence (WaLuDi) [227], can also be used in band selection tasks. The five bands selected using WaLuMI band selection are shown in Figure 3.26.

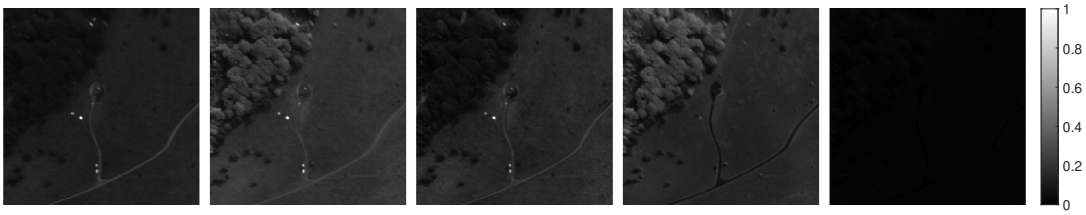


Figure 3.26: The five selected bands using WaLuMI Band Selection corresponding to bands $b = \{15, 31, 47, 76, 100\}$ with wavelengths $\lambda = \{476, 566, 658, 828, 971 \text{ nm}\}$

3.5.2 Spatial Dimensionality Reduction using Remote Sensing Indices

By investigating the spectral properties of the scene, spatial areas of interest can be selected and areas of non-interest can be discarded from further processing, often saving on large computational costs. Various measures can be used in such a way and are computed using ratios of wavelengths to create auxiliary features for any detection or classification tasks. Common measures in remote sensing applications include the NDVI [228], NDWI [115], and NDSI [116], as

well as their variants [229], which are routinely employed to estimate the amount of vegetation, water and snow present in a scene respectively. VIs such as NDVI and its variants are of particular interest in TD applications as they offer a simple and effective method to discriminate between vegetative and non-vegetative pixels. Three NDVI variants were selected and tested in discriminating between the desired background of vegetation and the foreground of synthetic materials to which the target objects of interest belong. Each of the methods used in this thesis are listed in Table 3.1.

Table 3.1: Vegetation indices used for spatial DR

Index	**Acronym**	**Equation**	**Reference**
Normalised Difference Vegetation Index	NDVI	$\frac{I_{\lambda_{NIR}} - I_{\lambda_{Red}}}{I_{\lambda_{NIR}} + I_{\lambda_{Red}}}$	Rouse et al. [228]
Normalised Difference Vegetation Index (red-edge)	NDVI_{re}	$\frac{I_{\lambda_{re}} - I_{\lambda_{Red}}}{I_{\lambda_{re}} + I_{\lambda_{Red}}}$	Hansen & Schjoerring [109] Ettehadi et al. [110]
Red-Edge Normalised Difference Vegetation Index	RENDVI	$\frac{I_{750\text{ nm}} - I_{705\text{ nm}}}{I_{750\text{ nm}} + I_{705\text{ nm}}}$	Gitelson & Merzlyak [111] Sims & Gamon [112]

An example of NDVI being applied to hyperspectral imagery is shown in Figure 3.27, where a vegetative scene from the UDRC Selene dataset (Appendix A.3.2) is processed. Figure 3.27a shows a pseudo-colour representation of the image, with Figures 3.27b and 3.27c showing the red and NIR bands used for obtaining the NDVI, shown in Figure 3.27d. The response of the vegetative pixels (red) is much higher in this image than other non-vegetative material (yellow).

NDVI and VIs in general were first used in remote sensing tasks which employed Landsat and similar early satellite-bourne imaging platforms. Chlorophyll, which serves as an indicator of vegetation health, absorbs visible light and reflects NIR light. By investigating the multiple reflectance values, generally in these red and NIR portions of the spectrum, VIs can be used as a proxy for the amount of chlorophyll present in each pixel and therefore as a measure of both the health and abundance of vegetation present in a scene [229]. As shown in Table 3.1, NDVI is computed as a ratio between intensities at wavelengths within these chlorophyll absorption and reflectance ranges and, as a result, this measure can be used to separate vegetative and non-vegetative pixels in an image, allowing for the masking of areas of non-interest.

The original definition of NDVI uses two image bands, one from each of the NIR and red parts of the spectrum, and is calculated using the first equation from Table 3.1. NDVI was initially developed for multispectral images, specifically the Multispectral Scanner (MSS) on Landsat 1, and as a result has a loose definition of NIR and red wavelengths with the respective Landsat 1 wideband images captured between $0.7 - 0.8\ \mu m$, $0.8 - 1.1\ \mu m$, and $0.6 - 0.7\ \mu m$

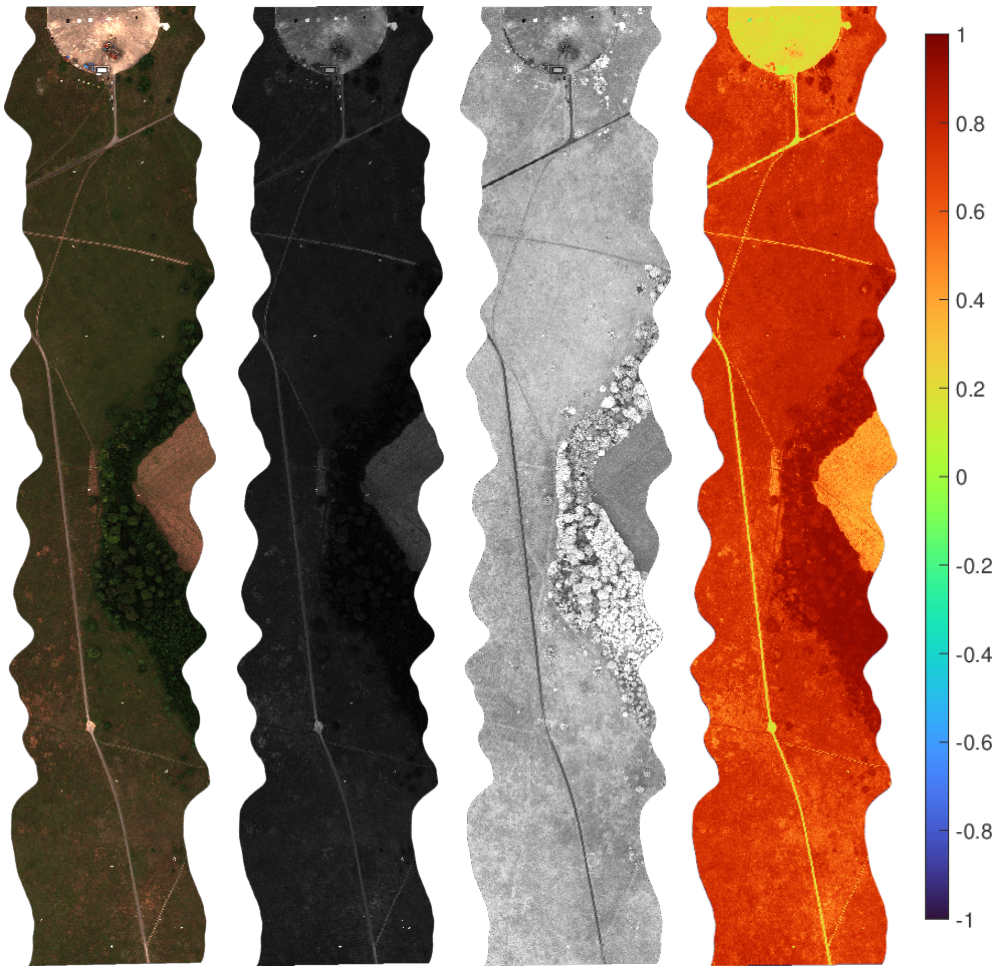


Figure 3.27: Example of NDVI on hyperspectral image data. **a)** False colour Hyperspectral Image. **b)** Red band ($\lambda = 650$ nm). **c)** NIR band ($\lambda = 810$ nm). **d)** False colour NDVI result.

for band 6 (NIR1), band 7 (NIR2), and band 5 (red) respectively. Alternatively, λ_{NIR} can be selected as a wavelength within the range of 760 nm to 860 nm and λ_{Red} as a wavelength within the range of 600 nm to 700 nm. Here $\lambda_{NIR} = 810$ nm and $\lambda_{Red} = 650$ nm are selected as they are close to the respective centre wavelengths of each range. NDVI represents the ratio of reflected red and NIR light at each pixel within the range -1.0 to 1.0 where large negative values indicate snow, water or cloud cover, values close to zero represent soil and rock where no vegetation is present and higher values represent higher concentrations of vegetative material.

An alternative definition, NDVI_{re} [110], changes the NIR wavelength of interest to one in the "red-edge", referring to the area between the red and NIR parts of the electromagnetic spectrum, where the rapid increase in reflection from vegetative material is visible. As most VIs are simple ratios between pixels it is possible, when using NDVI, to obtain similar ratios between

vegetation and certain synthetic materials. By instead investigating the red-edge it is possible to better distinguish between vegetative and non-vegetative pixels as it is almost exclusively vegetation which exhibits this particular red-edge response. The definition of NDVI_{re} is included as the second equation in Table 3.1 and a wavelength of $\lambda_{re} = 705$ nm was selected based on the Sentinel-2A band 5 centre wavelength, as used by the authors of [110].

Both NDVI and NDVI_{re} are wideband greenness VIs meaning they can be calculated from a single wavelength within an acceptable range or from multispectral bands which exhibit a significantly lower spectral resolution and wider sensor FWHM value. A third, narrowband, variant of NDVI, RENDVI, can also be defined. Here narrowband pertains to the definition of more precise wavelengths when calculating the RENDVI, commonly utilised with access to hyperspectral data. The narrowband RENDVI is calculated as a ratio between the red-edge and NIR responses using the third equation in Table 3.1 where a red-edge wavelength of 705 nm and a NIR wavelength of 750 nm are used. The RENDVI seeks to use the red-edge as well as NIR reflectance values as a measure of plant health which is correlated with the intensity of NIR light reflected in healthier leaves due to the higher chlorophyll levels present [229]. Using the VI values, the image can be segmented into areas of vegetative and non-vegetative pixels, under the assumption that targets are exclusively located in a subset of these non-vegetative pixels.

3.6 Target and Anomaly Detection Schemes

In this section, the five common classical methods for TD and AD investigated in this thesis are detailed. Four of these five algorithms, the Adaptive Coherence/Cosine Estimator, Constrained Energy Minimisation, the Spectral Angle Mapper and Spectral Information Divergence, are classified as TD algorithms. As such, they require *a priori* information about the targets to be detected in an image provided in the form of a reference or ground truth spectrum. The final method however, the Reed-Xiaoli Detector, does not require prior information about a target and finds outlying or anomalous pixels within the image. Whilst other TD algorithms such as Orthogonal Subspace Projection [117] are often used to good effect [145, 230], such methods require additional prior knowledge of the background which may not be fully known and as a result hinder the performance in a TD application [130, 231, 232] hence they are excluded from any analysis in later chapters.

The general form of a TD scheme can be represented as a pair of mutually exclusive hypotheses: the null hypothesis, where no target is present, denoted as H_0; and an alternative

hypothesis where a target is present with some non-zero abundance, a, denoted by H_1. This is expressed as

$$\begin{cases} H_0 : & \boldsymbol{x} = \boldsymbol{b} & \text{no target present} \\ H_1 : & \boldsymbol{x} = a\boldsymbol{t} + (1-a)\boldsymbol{b} & \text{target present} \end{cases} \tag{3.53}$$

where $\boldsymbol{x}$ is the observed spectrum of the PUT, $\boldsymbol{b}$ and $\boldsymbol{t}$ are the background and target spectra respectively, and a is some mixing coefficient, such that $0 \leq a \leq 1$.

3.6.1 Spectral Angle Mapper

The SAM [139] is a simple method which calculates the angle between the observed spectral signature and a target signature. The SAM algorithm follows the following hypotheses

$$\begin{cases} H_0 & \boldsymbol{x} = \boldsymbol{b} \\ H_1 & \mathbf{x} = \alpha\boldsymbol{t} + \boldsymbol{b} \end{cases} \tag{3.54}$$

where α is some variable accounting for the uncertainty in signal strength. The angle between two vectors, θ, is calculated as

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{t} \bullet \boldsymbol{x}}{\|\boldsymbol{t}\| \, \|\boldsymbol{x}\|}\right) \tag{3.55}$$

and the SAM detector can be implemented as

$$\text{SAM}(\boldsymbol{x}) = \frac{\boldsymbol{t}^T\boldsymbol{x}}{(\boldsymbol{t}^T\boldsymbol{t})^{1/2}\,(\boldsymbol{x}^T\boldsymbol{x})^{1/2}}. \tag{3.56}$$

where $\boldsymbol{t}^T\boldsymbol{x}$ is a measure of how aligned $\boldsymbol{x}$ and $\boldsymbol{t}$ are, and is normalised by the dividing by the product of the magnitudes, or lengths, of both $\boldsymbol{x}$ and $\boldsymbol{t}$.

3.6.2 Adaptive Cosine (Coherence) Estimator

The hypothesis of the ACE detector is given as

$$\begin{cases} H_0 & \boldsymbol{x} = \beta\boldsymbol{b} \\ H_1 & \boldsymbol{x} = \alpha\boldsymbol{t} + \beta\boldsymbol{b} \end{cases} \tag{3.57}$$

While similar to that used to derive the SAM detector from Equation (3.54), the hypothesis in Equation (3.57) contains an additional variable accounting for the uncertainty in strength of

the background, β. This better compensates for cases where the target forms only a component of the PUT, *i.e.,* in sub-pixel detection [123]. The ACE detection statistics are calculated using

$$\mathrm{ACE}(\boldsymbol{x}) = \frac{\left[(\boldsymbol{t}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]^2}{\left[(\boldsymbol{t}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{t}-\boldsymbol{\mu})\right]\left[(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]}, \tag{3.58}$$

where $\boldsymbol{x}$ and $\boldsymbol{t}$ are the observed and target spectra respectively, $\boldsymbol{\Sigma}$ is the sample covariance matrix, and $\boldsymbol{\mu}$ is the mean signature present in an image $\mathbf{X}$. The numerator is a measure of similarity between $\boldsymbol{x}$ and $\boldsymbol{t}$, adjusted by using the image covariance as an estimate of the background. Equation (3.58) is equivalent to measuring the cosine of the angle between $\boldsymbol{x}$ and $\boldsymbol{t}$ in the background whitened space.

3.6.3 Constrained Energy Minimisation

As discussed in Section 2.4, the CEM algorithm creates an optimal FIR filter based on the desired target spectrum whilst simultaneously minimising the energy of the background. The optimal FIR filter weights, $\boldsymbol{w}_{\mathrm{CEM}}$, are calculated using Equation (3.59) and the CEM TD statistic is calculated by multiplying the observed spectra by these weights as in Equation (3.60).

$$\boldsymbol{w}_{\mathrm{CEM}} = \frac{\boldsymbol{t}^T\boldsymbol{\Gamma}^{-1}}{\boldsymbol{t}^T\boldsymbol{\Gamma}^{-1}\boldsymbol{t}} \tag{3.59}$$

$$\mathrm{CEM}(\boldsymbol{x}) = \boldsymbol{w}_{\mathrm{CEM}}\boldsymbol{x} \tag{3.60}$$

In Equation (3.59), $\boldsymbol{\Gamma}$ is the sample correlation matrix, $\boldsymbol{\Gamma} = {}^1\!/_N\boldsymbol{X}\boldsymbol{X}^T$ where $\boldsymbol{X}$ has not been mean-centred. If the spectra are mean centred, the CEM algorithm becomes the Matched Filter algorithm [145].

3.6.4 Spectral Information Divergence

As discussed in Section 2.4, SID [140] is a measure of similarity between two signatures obtained by measuring the KL divergence, or relative entropy, between them. The SID is calculated as

$$\mathrm{SID}(\boldsymbol{x}) = D(\boldsymbol{x}\|\boldsymbol{t}) + D(\boldsymbol{t}\|\boldsymbol{x}), \tag{3.61}$$

where $D(\boldsymbol{x}\|\boldsymbol{t})$ is the KL divergence between $\boldsymbol{x}$ and $\boldsymbol{t}$, where, by obtaining the probability distributions of $\boldsymbol{x}$, and $\boldsymbol{y}$, $\boldsymbol{p}$ and $\boldsymbol{q}$, the KL divergence between $\boldsymbol{x}$ and $\boldsymbol{t}$, $D(\boldsymbol{x}\|\boldsymbol{t})$, is obtained by $\sum_{l=1}^{L} p_l \log(p_l/q_l)$. It can be noted that $D(\boldsymbol{x}\|\boldsymbol{t}) \neq D(\boldsymbol{t}\|\boldsymbol{x})$.

3.6.5 Global Reed-Xiaoli Detector

Global RXD is a popular algorithm used for AD [151] and is calculated as

$$\mathrm{RXD}(\boldsymbol{x}) = (\boldsymbol{x} - \mu)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \mu). \tag{3.62}$$

The RXD is equivalent to a squared Mahalanobis distance [153], taking advantage of the assumption that the image is mostly composed of some background distribution with targets being sparsely populated throughout. The RXD is a measure of distance from each pixel from the statistical background, such that outliers can be positively identified.

3.6.6 Orthogonal Subspace Projection

The OSP algorithm is based on the following hypotheses,

$$\begin{cases} H_0 & \boldsymbol{x} = \boldsymbol{B}\beta + \boldsymbol{n} \\ H_1 & \boldsymbol{x} = \alpha \boldsymbol{t} + \boldsymbol{B}\beta + \boldsymbol{n} \end{cases} \tag{3.63}$$

where α represents the uncertainty in the target strength, $\boldsymbol{B}$ represents the background subspace, β is the background coefficient basis vector, and $\boldsymbol{n}$ is a residual noise component. The OSP detector is derived in [117] as follows,

$$\mathrm{OSP}(\boldsymbol{x}) = \boldsymbol{t}^T \boldsymbol{P}_B^{\perp} \boldsymbol{x}. \tag{3.64}$$

where $\boldsymbol{P}_B^{\perp}$ is the background subspace calculated from the background Endmembers. This subspace projection is similar to the design of the CEM operator, however, in place of the *a posteriori* information provided in $\boldsymbol{\Gamma}^{-1}$, *a priori* knowledge of the background is required [145]. Due to this need to estimate the background which may not always be possible or practical for TD applications, the OSP detector is excluded from further analysis.

3.6.7 Performance Measures

In order to asses the performance of each of the TD algorithms, a number of measures are used in this thesis. Each of the various TD and AD algorithms investigated return a score representing the likelihood of the target spectrum appearing in each pixel. By varying the threshold above which a pixel is classified as a target, the various behaviours and performance of a TD algorithm can be assessed. Both Receiver Operator Characteristic (ROC) curves [233, 234]

and Precision-Recall (PR) curves [235, 236] are useful measures in determining an optimal operational threshold in order to maintain an acceptable False Alarm Rate (FAR). The Area Under the Curve (AUC) is a useful measure for comparing the ROC and PR behaviours of various algorithms. The ROC curve can be created by plotting the Probability of Detection (P_d), against the Probability of False Alarm (P_{fa}), at a series of thresholds.

Although ROC curves are a simple and effective way of rapidly visualising the performance of a classifier, it has been shown that ROC analysis can be flawed for imbalanced classes, as is the case for TD applications where many negative examples are shown with comparatively fewer positive examples. In [235], it is shown that PR curves are more informative for imbalanced classes as they correctly evaluate the fraction of True Positive (TP) detections amongst the total number of positive predictions, or the precision of the classifier. Precision can be calculated, using the number of FP detections, as $^{TP}/_{(TP+FP)}$. Recall is defined, using the number of False Negative (FN) predictions, as $^{TP}/_{(TP+FN)}$. Recall is the fraction of TP detections amongst the total number of positive examples. For TD applications, the number of positive examples is the total number of target pixels present in an image and PR curves can be obtained by plotting the precision of a classifier against its recall at a series of thresholds.

3.6.7.1 Statistical Measures

In addition to ROC and PR curve analysis, four other methods of assessing each of the target detection algorithms were considered. Three measures common in the assessment of binary classifier performance, the F1 score [237], Matthews Correlation Coefficient (MCC) [238], and Balanced Accuracy [239] were used. As the results from each TD algorithm can be evaluated as a binary classification between a target and the background, these measures provide insight into the performance of each TD algorithm.

The first metric used in analysing the performance of each of the detection algorithms is the F1 score. The F1 score is a common measure used in machine learning and statistics and can be calculated as

$$F_1 = \frac{2TP}{2TP + FP + FN}. \tag{3.65}$$

The F1 score has a value between 0 and 1. A score of 1 indicates perfect separation between target and background, whereas a score of 0 indicates that none of the target pixels were correctly identified. The F1 score can be described as the harmonic mean of precision and recall which is useful in applications where the precision and recall are equally important [240]. This is not always the case, for example in TD or medical imaging it is often important to limit the False Negative Rate (FNR). Therefore, placing an emphasis on the recall so as to maximise

the percentage of total TP detections can, in such applications, be advantageous.

The F1 score, while useful in many areas, does not consider the number of True Negative (TN) predictions *i.e.,* the number of pixels correctly not detected as a target. The MCC is another useful measure in determining the performance of a binary classifier, especially when the classes are imbalanced - like in anomaly or target detection where targets are sparsely populated in an image. The MCC can be calculated as

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \tag{3.66}$$

The MCC can be any value within a range of -1 to +1: with +1 indicating a perfect prediction; 0 indicating the classifier is equivalent to a random prediction; and a value of -1 indicating that the prediction and ground truths are inverted. As the MCC takes into account all four quadrants of the confusion matrix, an MCC score can be considered more informative than the corresponding F1 score as the MCC is only high when favourable results are obtained in each of the confusion matrix categories [241].

Accuracy is often another useful measure to consider when evaluating a classification task. However, a standard accuracy measure is inappropriate in cases where the classes are imbalanced. For example, in an imbalanced dataset, an event with probability 1e-6 could be misclassified (*i.e.,* TN = 1e6 and FN = 1) and the accuracy would still be in excess of 99.99%. Balanced Accuracy is an alternative measure which seeks to better represent the correct classification of negative examples as well as positive ones. Balanced accuracy can be calculated as

$$\text{Balanced Accuracy} = \frac{\text{sensitivity} + \text{specificity}}{2} = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right). \tag{3.67}$$

Measuring the balanced accuracy of the extreme example given previously, an accuracy of only 50% would be obtained, as all of the negative examples were correctly classified but the positive example was missed. In highly imbalanced classification problems the balanced accuracy is favourable as it apportions equal weighting to the positive and negative classifications.

3.6.7.2 Target Visibility

The fourth and final performance metric, the visibility [120], is a measure of how distinct the average target score is from the average background score. This is useful in assessing how the detection can be affected by applying DR and background suppression to the image data.

Visibility is calculated as

$$\text{Vis.} = \frac{\left|\bar{T}_t - \bar{T}_b\right|}{\max(T) - \min(T)}, \tag{3.68}$$

where, $\bar{T}_t$ and $\bar{T}_b$ are the mean scores of the target and background respectively, and $\max(T)$ and $\min(T)$ are the maximum and minimum score of the target in a scene at the output of any given detection algorithm. The visibility ranges from 0 to 1 with higher visibilities indicating more separation between a target and the background, in turn implying a higher probability of detection.

3.7 Compressive Sensing

3.7.1 Compressed Sensing Principles

CS works on two principles; sparsity, and incoherence [158, 242]. Sparsity is a property shared in all natural images and is the result of correlation in both spatial and spectral information [184]. The information contained in some signal or image, $\boldsymbol{x} \in \mathbb{R}^n$, can be represented in some basis $\boldsymbol{\Psi}$, where many of the elements in a vector of coefficients, $\boldsymbol{s}$, are small with the largest K being those that characterise the data, being so-called K-sparse,

$$\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{s}. \tag{3.69}$$

As a result, rather than measuring $\boldsymbol{x}$ directly, a random measurement $\boldsymbol{y} \in \mathbb{R}^p$, where $K < p << n$, can be used to capture the key information in $\boldsymbol{x}$ using some measurement matrix, $\boldsymbol{\Phi} \in \mathbb{R}^{p \times n}$,

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{s} = \boldsymbol{\Theta}\boldsymbol{s}. \tag{3.70}$$

Incoherence refers to the structure of the measurement matrix, $\boldsymbol{\Phi}$, in relation to the basis matrix $\boldsymbol{\Psi}$. If these two matrices are coherent, *i.e.,* their elements are correlated, the columns in the resultant sensing matrix, $\boldsymbol{\Theta}$, will be similar, or in the worst case identical. If these columns are identical, there is no way to identify which of them contributed to the final measurement and therefore determine the sparse components of $\boldsymbol{s}$. The incoherence principle dictates that $\boldsymbol{\Phi}$, unlike the signal of interest, has a dense representation in $\boldsymbol{\Psi}$. This is desirable as it effectively ensures sufficient, diverse, information is captured from the sparse set of coefficients in $\boldsymbol{s}$ such that an acceptable reconstruction can be achieved [158, 184]. As a result, the design of the measurement matrix in any CS scheme is key. An example of this is shown pictorially in Figure 3.28 where $\boldsymbol{\Phi}$ is a random Gaussian matrix and $\boldsymbol{\Psi}$ is the DCT basis matrix.

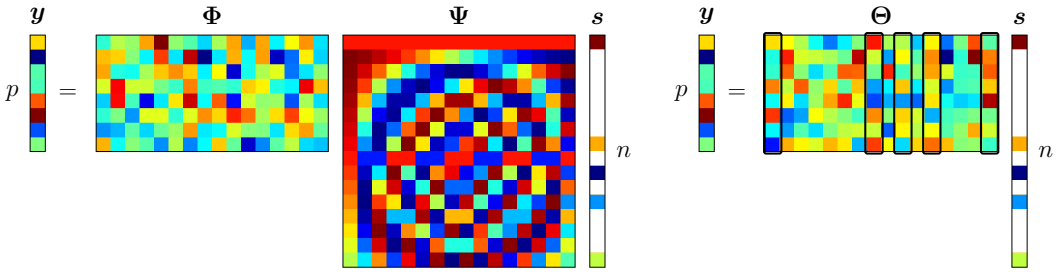


Figure 3.28: Example of the CS measurement process adapted from [158]. **a)** CS Measurement of a $k = 5$ sparse signal where $\boldsymbol{\Phi}$ is a random Gaussian matrix and $\boldsymbol{\Psi}$ is the DCT basis. **b)** Measurement process using $\boldsymbol{\Theta}$.

In order to recover $\boldsymbol{s}$ from the measurement $\boldsymbol{y}$, the underdetermined system presented in Equation (3.70) must be solved. This is achieved by finding a solution which minimises the l_1 norm of $\boldsymbol{s}$, $\|\boldsymbol{s}\|_1$, promoting the sparsity of $\boldsymbol{s}$ in any solution. This minimisation is given as

$$\hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} \|\boldsymbol{s}\|_1 \ \ \textbf{st.} \ \ \boldsymbol{\Theta s} = \boldsymbol{y}. \tag{3.71}$$

With this estimate of the sparse coefficients vector, $\hat{\boldsymbol{s}}$, an estimate of the original data can be obtained using $\boldsymbol{\Psi}$,

$$\hat{\boldsymbol{x}} = \boldsymbol{\Psi}\hat{\boldsymbol{s}} \tag{3.72}$$

3.7.2 Coded Aperture Imaging

CSI, and of particular relevance to the contributions in Chapter 6 of this thesis CA-based imaging, offers a practical, hardware, implementation of the principles outlined in Section 3.7.1. CASSI, and its derivative architectures, uses a CA in order to encode spatial and spectral information into a single compressed measurement. The CA should be designed such that it promotes incoherence with $\boldsymbol{\Psi}$ in order to obtain sufficient information for reconstruction [193]. Additionally, dispersive prisms are used to laterally shear spectral information to be incident at different spatial locations in order to effectively encode a 3D spectral image in a 2D coded measurement. The behaviour of some CA, as well as the intrinsic optical elements within a CASSI system, is described in the matrix, $\boldsymbol{H}$, such that a compressed measurement, $\boldsymbol{y}$, of some input image, $\boldsymbol{f}$, is given as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{f}. \tag{3.73}$$

Assuming that $\boldsymbol{f}$ is sparse on some basis, $\boldsymbol{\Psi}$, Equation (3.73), can be rewritten as $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{\Psi}\boldsymbol{s}$, and is identical to Equation (3.70), where following the notation in CASSI literature [170, 184], $\boldsymbol{\Phi}$ is replaced with the sensing matrix defined by some CA, $\boldsymbol{H}$, (discussed further in Chapter 6) and the effective sensing matrix, $\boldsymbol{\Theta}$, becomes $\boldsymbol{A}$. With these substitutions, a single CASSI measurement is obtained using $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{s} = \boldsymbol{H}\boldsymbol{\Psi}\boldsymbol{s}$.

Similarly to the general method for reconstructing the original signal from a compressed measurement presented in Equation (3.71), CASSI measurements can be reconstructed via optimisation. The general form of this optimisation is given as

$$\hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{s}\|_2^2 + \lambda \|\boldsymbol{s}\|_1 , \tag{3.74}$$

where the constraint that the produced solution must exactly match the observed measurement, *i.e.,* $\boldsymbol{A}\boldsymbol{s} = \boldsymbol{y}$ is replaced with an l_2 norm-based data fidelity term. This is coupled with the l_1 norm term using a regularisation constant, λ, allowing for a trade off between the promotion of a sparse solution while also ensuring an accurate reconstruction [184, 193]. Once an approximation of the sparse coefficient vector, $\boldsymbol{s}$, has been obtained, the original scene can be reconstructed, substituting $\hat{\boldsymbol{x}}$ for $\hat{\boldsymbol{f}}$ in Equation (3.72).

3.7.2.1 Single Dispersion CASSI (SD-CASSI)

The first CSI model investigated in this thesis is SD-CASSI, it consists of three main elements; the CA which encodes spatial information, a dispersive element which distributes spectral information over some spatial extent, and a focal plane array to capture the encoded and shifted image. Figure 3.29 shows the optical path through an SD-CASSI system.

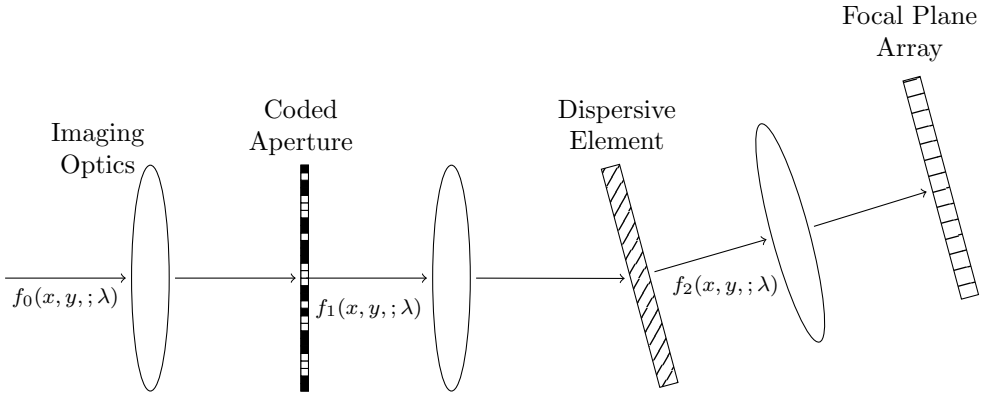


Figure 3.29: Optical flow through a SD-CASSI system.

From Figure 3.29, the spectral density of light signals entering the optics is assumed to be $f_0(x, y; \lambda)$, where x and y are points in the two spatial components of the image and λ is the

spectral component at the point (x, y). Any CA with a discrete pattern $T(x, y)$, with elements $t \in 0, 1$, represents the transmission of light through the aperture. The spectral density of incoming light can be encoded by T using;

$$f_1(x, y; \lambda) = f_0(x, y; \lambda)T(x, y). \tag{3.75}$$

The resultant coded field is sheared by a dispersive prism which modulates this encoded signal in the spectral dimension. The spectral density at this stage, $f_2(x, y; \lambda)$, can be realised using;

$$f_2(x, y; \lambda) = f_0(x + S(\lambda), y; \lambda)T(x + S(\lambda), y). \tag{3.76}$$

For the purposes of the simulations presented in Chapter 6, the dispersion of the prism, $S(\lambda)$, is assumed to be linear over the range of wavelengths considered. The image incident on the detector, $g(x, y)$ is

$$g(x, y) = \int f_2(x, y; \lambda)d\lambda, \tag{3.77}$$

where the light incident on the detector is the integrated, spatially shifted, spectral channels. This distributes the spectral information across a linear region, corresponding to the number of observed spectral components, along the direction of dispersion of the prism. This gives the appearance of blurring or smearing of objects within the imaged scene as the spectral information and energy relating to a single pixel is spread across some spatial extent. A simplified representation of the encoding and dispersion performed using SD-CASSI is shown in Figure 3.30.

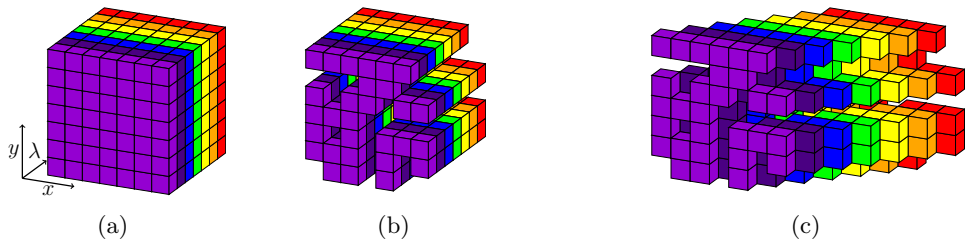


Figure 3.30: Sampling of a Multispectral Image using SD-CASSI. **a)** Input datacube. **b)** Spatio-spectral encoding. **c)** Spatio-spectral dispersion.

Figure 3.30a displays a spectral datacube similar to those presented in Section 3.4.1. Figure 3.30b shows the datacube at position $f_1(x, y; \lambda)$ in Figure 3.29 where the image has

been punctured by some binary CA. Finally, Figure 3.30c shows the effect of the prism on the encoded datacube at position $f_2(x, y; \lambda)$, where each band is sheared as a function of its wavelength with the spatial information in each pixel being spread across multiple locations. This spread datacube is then integrated in the λ axis to form the SD-CASSI measurement.

3.7.2.2 Dual Dispersion CASSI (DD-CASSI)

A similar optical flow can be derived for the DD-CASSI system [243]. In a DD-CASSI system, two dispersive elements are used, each dispersing incoming light in opposite directions, $\pm S(\lambda)$. This has the effect of first shearing the observed scene, before encoding by some CA, T, followed by a reversal in the shearing operation. This second shearing operation returns the now encoded image back to the original domain which negates the smearing effect encountered using an SD-CASSI system. Figure 3.31 shows the optical path through a DD-CASSI system.

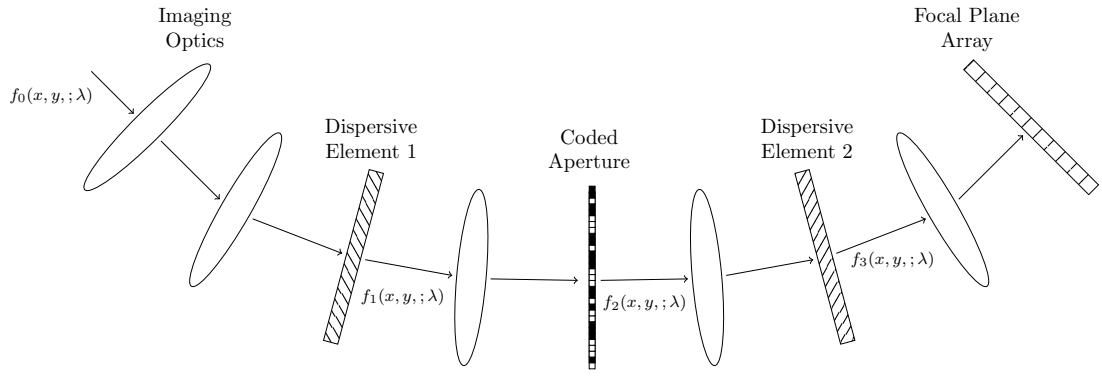


Figure 3.31: Optical flow through a DD-CASSI system.

The spectral density entering the instrument is, again, denoted as $f_0(x, y; \lambda)$. The density after the first dispersive element, $f_1(x, y; \lambda)$, is a function of the dispersion of the first prism, $S(\lambda)$;

$$f_1(x, y; \lambda) = f_0(x + S(\lambda), y; \lambda), \tag{3.78}$$

After being dispersed by the first prism, the incoming light is then encoded by a CA, $T(x, y)$

$$f_2(x, y; \lambda) = T(x, y) f_0(x + S(\lambda), y; \lambda), \tag{3.79}$$

where the CA is required to be larger in comparison to an SD-CASSI system to account for the spread caused by the first dispersive element. It can be noted that this CA provides both spatial and spectral encoding as, unlike the SD-CASSI model, the image is sheared prior to encoding.

This dispersion is reversed by the second dispersive element and $f_3(x, y; \lambda)$ is described prior to the detector as

$$f_3(x, y; \lambda) = f_0(x, y; \lambda)T(x - S(\lambda), y). \tag{3.80}$$

Finally, $f_3(x, y; \lambda)$ is integrated with respect to λ at the focal plane array to form the compressed measurement $g(x, y)$,

$$g(x, y) = \int f_3(x, y; \lambda)d\lambda \tag{3.81}$$

where, near-identically to the SD-CASSI model, the light incident on the detector is the integrated output from the second dispersive element. This second dispersive element relocates the dispersed, encoded spectral data back into the original domain of $f_0(x, y; \lambda)$. This avoids the blurring present in SD-CASSI and concentrates the energy from a single pixel in the input to a single pixel of the observed measurement. A simplified representation of the encoding and dispersion performed using DD-CASSI is shown in Figure 3.32.

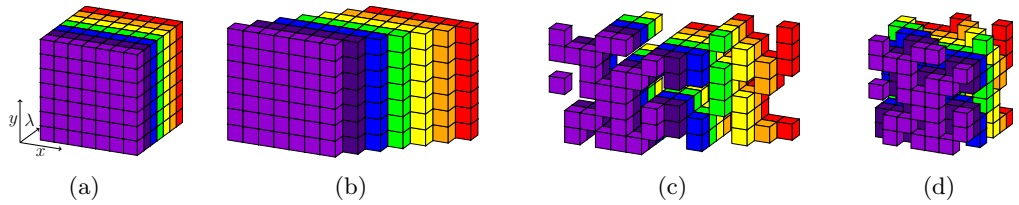


Figure 3.32: Sampling of a Multispectral Image using DD-CASSI. **a)** Input datacube. **b)** Spatio-spectral dispersion. **c)** Spatio-spectral encoding. **d)** Reverse dispersion.

Figure 3.32a shows the same datacube present in the example for SD-CASSI in Figure 3.30a. This datacube, however, is sheared prior to encoding, as shown in Figure 3.31 and can be visualised as shown in Figure 3.32b, representative of the status of the datacube at position $f_1(x, y; \lambda)$. This dispersed datacube is then encoded by some CA, required to be larger than the equivalent SD-CASSI CAs due to the initial shear. This has an identical effect to the encoding from Figure 3.30b, however, the puncturing affects data from more than one spatial location. Pixels removed in certain spatio-spectral locations, will be sufficiently represented by spatially and spectrally adjacent pixels in the measurement. Finally, the shearing is reversed as shown in Figure 3.32d and the DD-CASSI measurement is made by integrating on the λ axis.

3.7.3 Coded Aperture Design

Intrinsic to any CASSI system is the CA used for modulating and filtering the light through the optical elements of the imager. Designing optimal CAs is critical in the reconstruction of

the compressed image incident on the sensor [193], ensuring sufficient information is retained to accurately represent the original scene. Even in applications where reconstruction is not performed, such as the proposed compressed TD scheme in Chapter 6, the design of CAs is an important consideration and multiple methods of CA generation are detailed in this section. Blue noise sampling has been shown to be useful in such CASSI approaches [184, 193, 244–246], and so is a focus of CA design here.

3.7.3.1 Random Coded Apertures

The first type of CAs considered are *random* CAs. Here, random refers to unstructured pseudo-randomly generated matrices which have been subsequently thresholded to form binary CAs, or Bernoulli matrices. In situations where an ensemble of CAs are generated, *i.e.,* where multiple shots are taken, the correlation between individual CAs remains uncontrolled, which can often lead to sub-optimal sampling. Two CA classes are drawn from such pseudo-randomly generated matrices, henceforth referred to as "Random" CAs and "Boolean" CAs.

A) Random Coded Apertures

   In this method, there are no constraints placed either spatially or temporally, where temporal resolution is dependent on the number of shots, K, being generated. As a result, pixels can be sampled multiple times over successive shots or, alternatively, can also be missed entirely which in scenarios where small or sparse detail is of interest is clearly undesirable.

B) Boolean Coded Apertures

   Boolean CAs are similar to Random CAs, however, they have an additional constraint that each pixel must be sampled once and only once. Such CAs do not suffer from oversampling (sampling a pixel multiple times over an ensemble) or undersampling (pixels being missed).

Both Random and Boolean CAs tend to have poor spatial distribution as there is no spatial regularisation for determining or limiting the clustering exhibited in the produced ensemble. It can be noted that, when $K = 1$, CAs produced using each of these methods have similar characteristics due to a maximum transmittance of 0.5, *i.e.,* a maximum of half the CA pixels being either zero or one.

3.7.3.2 Poisson Disc Coded Apertures

The fast Poisson Disc (PD) sampling algorithm, described in [247], was used to generate CAs in Chapter 6. Each CA in the ensemble was created independently by iteratively generating n

random points with some minimum spacing, r, between them. As a result, these CAs produce a spatial blue noise distribution, exhibiting a slight suppression of low frequencies. Similarly to a Random CA, an ensemble of PD CAs have the potential to oversample and undersample pixels, as this approach equates to a dart throwing algorithm. While it can generate samples in N-dimensions, *i.e.,* an ensemble of CAs in a single pass, this produced excessively sparse CAs in practice. Another similar algorithm, although not considered here, is Mitchell's Best Candidate which has been used for blue noise sample creation [248].

3.7.3.3 Void-and-Cluster Blue Noise Coded Apertures

The Void and Cluster (VAC) algorithm [249] is a technique for performing halftoning, *i.e.,* the rendering of binary images which emulate a greyscale image, via the creation of an optimally spaced dithering matrix. This dithering matrix can then be used to generate an ensemble of CAs. The VAC algorithm consists of four steps as described in the original paper for generating an $M \times N$ dither array;

1. A randomly generated binary pattern is uniformly distributed using the method shown in Figure 3.33, the number of ones present in this image is stored as the "*Rank*" of this image.

2. A "minority" pixel, in this case a 1, is removed from the centre of the tightest cluster of ones, *i.e.,* that with the most surrounding pixels, and in a separate dither array, the current Rank value is inserted into the corresponding location. The Rank value then decreases by one and this is repeated until the Rank value reaches 0.

3. The original pattern is then recovered and the algorithm recursively finds the location of the largest void, or cluster of zeroes, and fills the pixel at the centre with a 1 and the Rank is inserted into the dither array. The Rank is then increased by one each iteration and the stopping condition is met at Rank $= (MN)/2$.

4. This phase is similar to the previous, with the exception being that the notion of a minority pixel is then changed from 1 to 0 as there are now more ones in the binary pattern. As such, clusters of zeroes are targeted and filled with ones until the Rank $= MN$. As a result, the binary pattern consists solely of ones, and the dither array contains every unique value between 0 and $MN - 1$.

Step 1 is displayed in Figure 3.33 steps 2 – 4 are displayed in Figure 3.34. The dither matrix forms the base of the VAC ensemble CA, and is normalised using the approach from the

original paper [249], quantising the array into K levels corresponding to the K shots taken with a multi-shot CASSI system. Because the numerically adjacent Rank values are dispersed optimally in the generated dither array, the resultant binary images, corresponding to the set of pixels where the quantised dither array equals the shot number, exhibit a strong blue-noise like distribution. If the number of CASSI shots, K, is equal to 1, the number of quantisation levels is set to 2 to ensure that the transmittance in the single CA is approximately 0.5.

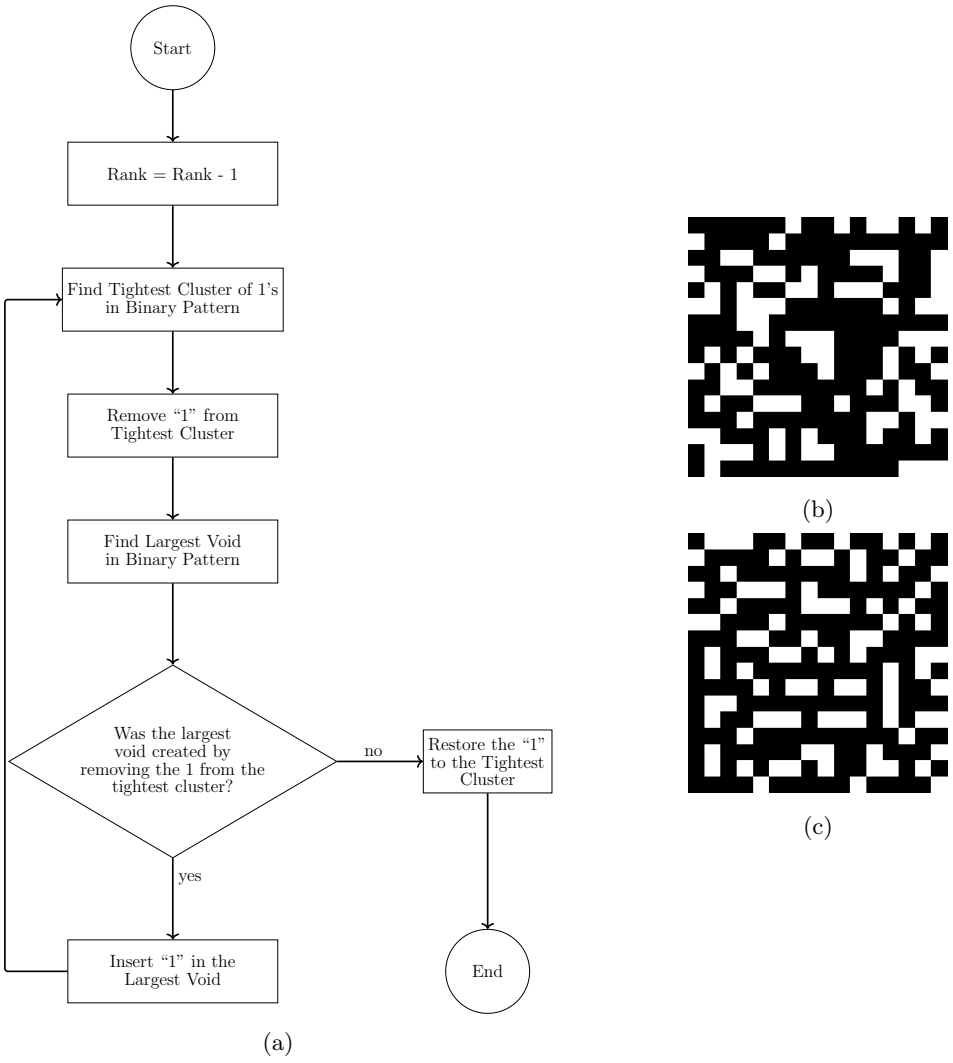


Figure 3.33: Initial Blue Noise Pattern Creation for Void-and-Cluster **a)** Flowchart showing the stages to create the initial binary pattern for input into the Void-and-Cluster algorithm to create blue noise Coded Apertures. **b)** Random binary image created using white noise. **c)** Uniformly distributed version of b) using the process displayed in a).

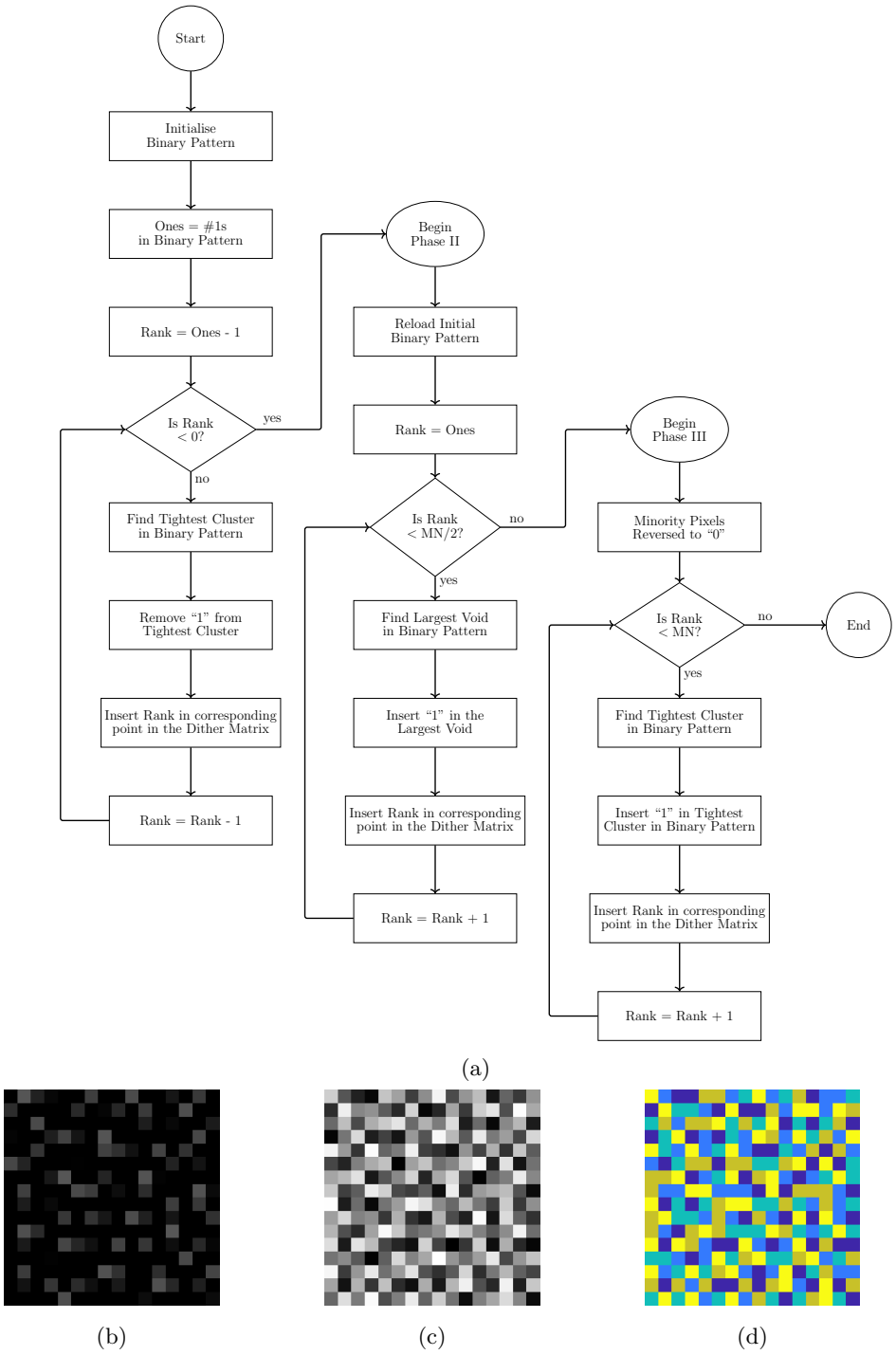


Figure 3.34: **a)** Flow chart showing the phases of the Void and Cluster Algorithm [249]. **b)** Dither array after phase 1 where the tightest clusters of Figure 3.33c have been filled. **c)** Dither array after phase 3 where the largest voids of Figure 3.33c have been filled. **d)** The dither array from c) quantised to $K = 5$ levels with false colour applied.

3.7.3.4 SpatioTemporal Blue Noise Coded Apertures

Although it produces high quality blue noise textures for use as CAs, the VAC algorithm can take a long time to converge, especially for large matrices. Other methods for generating blue noise such as the PD algorithm [247] or Mitchell's Best Candidate [250], offer a good approximation of blue noise samples but do not produce a blue noise mask, or texture, as the VAC algorithm does. The ability to generate a blue noise mask, allows for each pixel in an image to be sampled, much like the distinction between Random and Boolean CAs. This is advantageous in that both oversampling and undersampling are avoided but with the added benefit of introducing some structure in the spatial distribution of sampling points.

An alternative to VAC for producing blue noise mask-based CAs is Spatiotemporal Blue Noise (STBN) [193]. STBN uses the assumption that if local regions within a CA exhibit blue noise characteristics, then so too will the CA as a whole. Local blue noise-like patterns are achieved through a set of simple design criteria applied to each pixel in a CA. These are; maximising horizontal, vertical, and diagonal spacing within each of the K shots based on some weighting in each direction, and limiting each spatial location to be sampled only once in each ensemble. This latter criterion removes any temporal redundancy and avoids both oversampling and undersampling. The STBN algorithm takes an initial Boolean CA as a guess and attempts to iteratively optimise a CA ensemble based on these criteria. However, compared to the VAC algorithm, the blue noise textures exhibit slightly weaker low frequency noise suppression due to the spacing between samples only being optimised in the cardinal and intercardinal directions.

3.7.4 Coded Aperture Generation

Each of the methods for generating CAs presented in Section 3.7.3 were used to generate the examples given in Figure 3.35 with spatial dimensions $M = N = 64$ and number of shots $K = 3$. From top to bottom, the rows of Figure 3.35 show the ensemble CAs, *i.e.,* where each of the $K = 3$ shots is shown as a channel of the resultant composite RGB image, meaning red pixels are sampled in the first shot, green in the second, and blue in the third. Any other colours indicate the pixel is sampled multiple times based on the RGB additive colour model, *i.e.,* cyan, magenta, and yellow indicate the pixel has been sampled in two shots, white indicating sampling in all three, and black pixels indicating no sampling has occurred. This is present in both the Random and PD CAs as there is no constraint on each pixel being sampled once. This row is followed by the binary CA generated for each of the three individual shots. Here, the disparity between the randomly generated CA types and the more structured blue noise

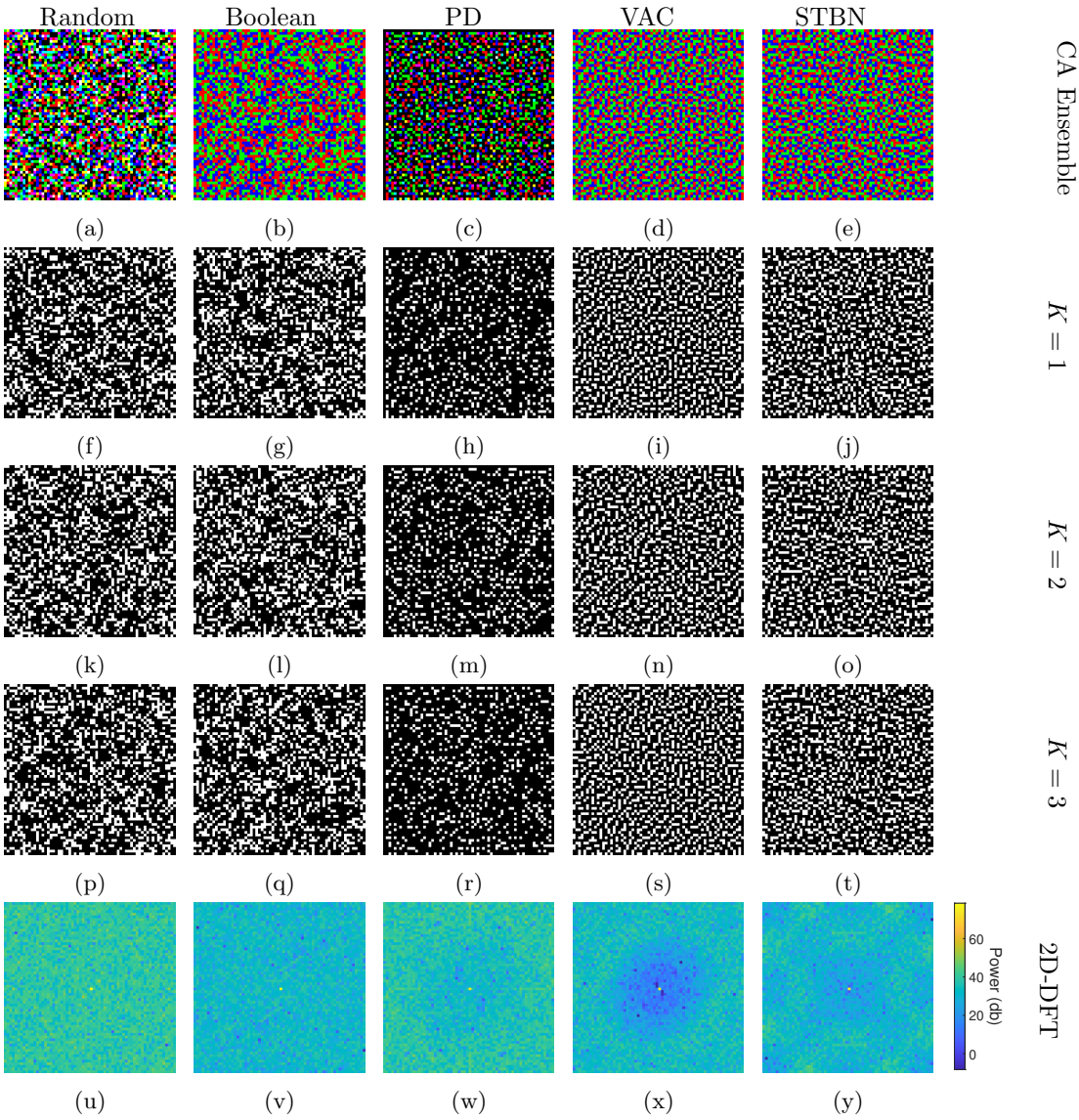


Figure 3.35: Examples of Ensemble Coded Apertures using each of the generation techniques. The columns from left to right show the random, Boolean, Poisson disk, void-and-cluster, and spatio-temporal blue noise methods of generation respectively. The rows from top to bottom show the generated false colour CA ensemble, shots at $K = 1$, $K = 2$, $K = 3$ and the result of taking the 2D-DFT to view the frequency density of each method.

CAs is visible. The final row displays the 2-Dimensional Discrete Fourier Transform (DFT) shifted such that the DC, or 0 Hz, component is in the centre of the image, where blue colours indicate low magnitude frequency components and yellow colours indicate high magnitude - normally reserved for the DC component. The value of the image moving outwards radially from the centre indicates the frequency density of the ensemble CA. These plots highlight the

low-frequency suppression characteristics of blue noise masks, *i.e.,* the PD, VAC, and STBN methods. By optimally spacing each individual sample, the resultant CA has minimal, or is devoid of, large areas of continuous foreground. Conversely, the Random and Boolean CAs have a more even distribution across the range of frequencies considered.

The CA designs described in the previous section range in their complexity and this is reflected in the methods for generating them. Figure 3.36 shows the time taken to generate each of the CA types considered in this thesis over a range of desired resolutions from $M = N = 2^2$ (*i.e.,* 4×4) to $M = N = 2^9$ (*i.e.,* 512×512). Each method was repeated five times for each of the resolutions considered and the average generation times of each method are displayed.

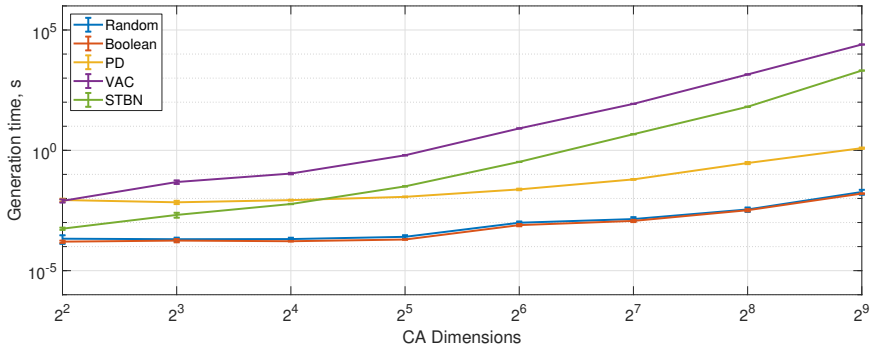


Figure 3.36: Coded Aperture generation algorithm run time.

The Random and Boolean CAs are generated consistently faster than the other methods, often by orders of magnitude. This is expected, as they are created using threshold operations on randomly generated matrices and are therefore the simplest of the generation methods considered, exhibiting slight rises in generation time with an increase in resolution. The PD method also increases slightly as the resolution increases, however, it takes an order of magnitude longer than the Random and Boolean methods. Due to its iterative nature, the number of points is also a factor in the generation time. The time taken to perform the VAC and STBN algorithms, however, increases exponentially as the resolution increases. This is a result of the number of operations required at each resolution doubling compared to the last. The average time taken to generate each CA at the set of desired resolutions is listed in Table 3.2.

Table 3.2: Average CA generation time for each method considered at increasing resolutions.

	2^2	2^3	2^4	2^5	2^6	2^7	2^8	2^9
Random	2.13×10^{-4}	2.03×10^{-4}	2.05×10^{-4}	2.55×10^{-4}	9.81×10^{-4}	1.39×10^{-3}	3.49×10^{-3}	1.86×10^{-2}
Boolean	1.60×10^{-4}	1.78×10^{-4}	1.67×10^{-4}	1.97×10^{-4}	7.80×10^{-4}	1.17×10^{-3}	3.25×10^{-3}	1.60×10^{-2}
PD	8.61×10^{-3}	6.92×10^{-3}	8.52×10^{-3}	1.17×10^{-2}	2.38×10^{-2}	6.19×10^{-2}	2.97×10^{-1}	1.23
VAC	7.94×10^{-3}	4.82×10^{-2}	1.08×10^{-1}	6.22×10^{-1}	8.12	8.53×10^{1}	1.42×10^{3}	2.49×10^{4}
STBN	5.55×10^{-4}	2.07×10^{-3}	5.87×10^{-3}	3.18×10^{-2}	3.31×10^{-1}	4.67	6.45×10^{1}	2.06×10^{3}

3.8 Deep Learning Approaches to Object Detection and Morphology

Another increasingly common method for object detection is the use of deep learning techniques. Deep learning has been employed to great effect in various image processing applications such as segmentation, edge detection as well as for object tracking [251] and object detection and subsequent classification [252]. Deep-learning based methods, especially CNNs, offer object detection typically based on some bounding box regression and classification [253].

CNNs are commonly used to classify images by detecting and classifying features of objects present in a scene. By training on large quantities of data, very deep features can be learned and trained models can be applied to unseen images to often good effect. Region-CNN, or Region Convolutional Neural Network (R-CNN), [254] was developed to improve on traditional feature detectors like the Scale-Invariant Feature Transform (SIFT) and other histogram based methods and seeks to apply the classification quality of CNNs to object detection. Rather than the objective being to classify the image as a whole, Region of Interests (ROIs) pertaining to discrete objects within an image are classified. These regions first need to be defined and this can be achieved using a region proposal method such as selective search or edgeboxes.

The R-CNN pipeline is shown in pictorial form in Figure 3.37. For each image, N region proposals are generated with each region proposal being reshaped and fed through a trained CNN, typically AlexNet, VGG-16 or ResNet 101. This produces a feature vector that can then be classified using a per-class SVM, producing a class for each of the proposals. At this stage, the regions can be refined by discarding those which overlap with others with a higher score. The bounding box from the object is also refined using a regression model so as to try and maximise the intersection over union (IoU) with the ground truth.

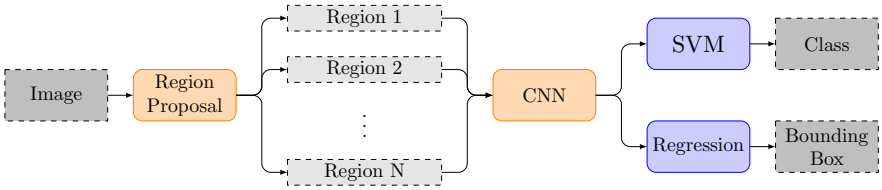


Figure 3.37: Flowchart detailing the stages in the R-CNN object detection framework.

Although R-CNN works very well, its efficacy in real time or fast applications is limited due to it requiring a forward pass of the CNN for every region proposal in every image. This results in excessive redundancy which leads to an overhead in run-time. Another limitation on the execution time is the requirement to train three separate models; the CNN for image

features, the SVM to classify the object; and the regression model to fit the bounding boxes. An improvement on this which sought to rectify these shortcomings was later developed as the Fast R-CNN [255].

Fast R-CNN seeks to improve on the two main drawbacks of R-CNN, namely the lengthy processing time by combining all three models into one which can train the CNN, classifier and bounding box regression. Another improvement is that of the region proposal stage, rather than each proposed region being input into the CNN, the region proposals are used on the feature map output from the CNN. As a result, this process takes place only once in a ROI pooling stage - this also deals with redundancy from overlapping bounding boxes. The Fast R-CNN pipeline can be seen in Figure 3.38.

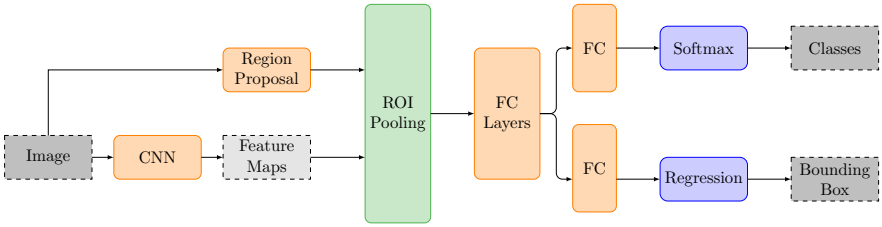


Figure 3.38: Flowchart detailing the stages in the Fast R-CNN object detection framework.

While Fast R-CNN offers an improvement on run-time, further improvements were made in Faster R-CNN [253] by including a region proposal network (RPN) that replaces the external region proposal methods (which were identified as the bottleneck of this previous process). The pipeline for Faster R-CNN object detection is shown in Figure 3.39. Faster R-CNN offers the fastest implementation of all three.

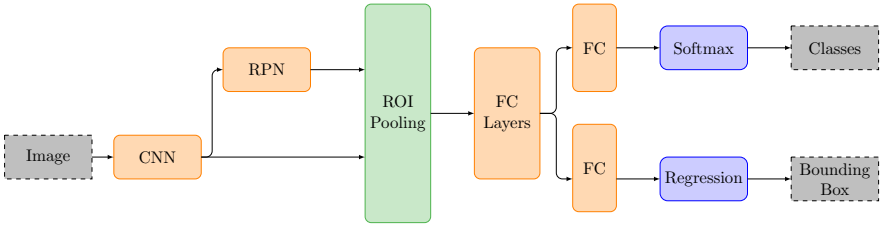


Figure 3.39: Flowchart detailing the stages in the Faster R-CNN object detection framework.

R-CNN, Fast R-CNN and Faster R-CNN, You Only Look Once (YOLO), the Single-Shot Detector (SSD) and other common Machine Learning methods for object detection are reviewed and compared in [256].

Morphology, unlike the bounding-box-based networks described, provides a rapid, pixel-wise decision for object detection and pattern recognition based on size and shape without the

need for training data. Recently, CNNs have been used to perform semantic segmentation or instance segmentation in order to isolate pixels belonging to some target object. Networks such as Mask R-CNN [257] offer similar pixel-wise classification as would be expected from more morphological methods. However, such deep learning techniques are comparatively complex and computationally expensive. Mask R-CNN is nearly identical to Faster R-CNN, with an additional parallel process for creating a mask based on the returned ROIs as shown in Figure 3.40.

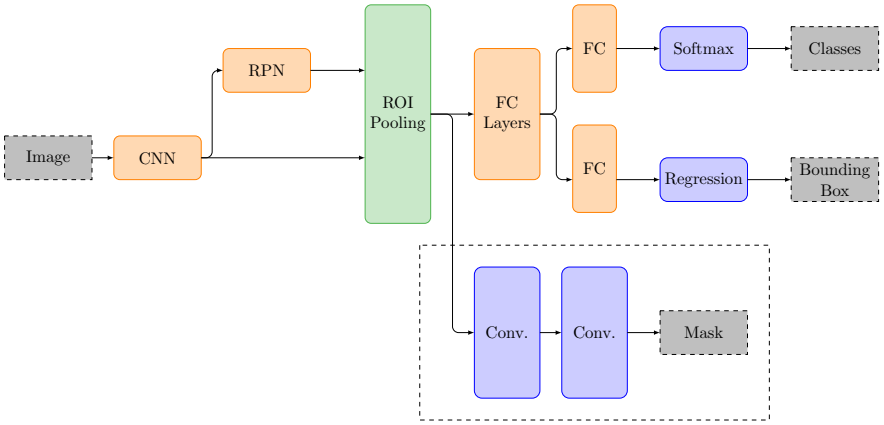


Figure 3.40: Flowchart detailing the stages in the Mask R-CNN object detection framework.

Recently, various investigations into ways of combining morphology into state of the art deep learning methods have been introduced. All morphological operations are performed using a structuring element, or SE, and as such the design of these SEs is of great importance. Traditionally, SEs are designed manually with some prior knowledge of the operation. However, more automatic methods to optimise the design of SEs using supervised [258, 259] and deep learning [260–262] methods have been explored. Morphology can also be deployed within deep learning frameworks, forming the non-linear filters required in feature pooling stage [263] or in other layers within a CNN [264]. Deep learning can also be used to learn and mimic morphological operations such as erosion and dilation [263] as well as the HMT [265] or used in morphological operations for various image analysis tasks [261, 266]. All of these proposed methods are solely restricted to single channel images, or can be applied marginally to colour images to provide a naïve multi-channel approach. In the future, investigation into how to couple the presented morphological techniques with deep learning may be taken, however the immediate aims of this work are to improve upon the field of colour morphology.

3.9 Summary

In this chapter, the technical background pre-requisite to the contributions presented in this thesis have been presented.

First, the origins, mathematical fundamentals, and extensions of Mathematical Morphology in binary and greyscale images have been explored. The need for, and difficulty behind, an extension to colour and multivariate images is discussed and existing methods for performing morphological operations on such images are displayed. This foundation is built upon in Chapter 4 where a novel multivariate Hit-or-Miss Transform based on reduced ordering is formalised and discussed.

Secondly, an overview of spectral imaging methods and their advantages, disadvantages, and uses are discussed. As these techniques generate such large quantities of data, a number of common Dimensionality Reduction techniques are derived for use in compressing such images. Additionally, as efficient Target Detection is of particular interest in this thesis, a number of benchmark TD algorithms are derived and discussed. Combining these techniques to improve the efficiency of spectral target detection is explored in Chapter 5 of this thesis.

An overview of the general Compressed Sensing problem is given along with how it relates to Compressive Spectral Imaging and particularly, the Coded Aperture Snapshot Spectral Imaging camera. Using the models derived in this chapter, along with the methods of CA generation discussed, investigations into performing TD in the compressed domain are explored further in Chapter 6.

Finally, the development of the Region-based Convolutional Neural Networks was discussed as this presents an example of a family of deep learning methods able to perform object detection and localisation in images. While not a focus of this thesis, deep learning-based techniques are prevalent in academic literature and may also benefit from, or augment, some of the signal and image processing approaches explored in this work.

Chapter 4

Extending the Morphological Hit-or-Miss Transform to Multivariate Images using a Reduced Ordering Scheme

4.1 Introduction

Mathematical Morphology is a powerful, yet elegantly simple, set of tools for image analysis. Despite its prevalence in image analysis applications on binary and greyscale, single channel images, there are no generally accepted implementations of MM for use in multi-channel images. As discussed in previous chapters, this is due to the lack of a natural ordering relationship in multivariate data, a pre-requisite for the definition of morphological operators.

In this chapter a new method for applying morphological techniques, specifically the morphological HMT, to colour and spectral multivariate data is presented. This method is based on a reduced ordering relationship for multivariate data combined with the noise-robustness offered by relaxed morphology using rank order filters. The novel morphological operation presented in this chapter, the Multi-Dimensional Percentage Occupancy Hit-or-Miss Transform (MDPOHMT), like other HMTs relies on two, non-intersecting, SEs; one which probes the foreground, $\boldsymbol{S}_{FG}$, and one which probes the background, $\boldsymbol{S}_{BG}$, in order to detect suitable objects of interest. By extending the spectral dimensionality of the SE to match that of the image under test, the similarity between corresponding image and pixels within the support of an SE can be derived using a reduced ordering. The reduced ordering is applied on a pixel-wise basis which allows for both flat and non-flat SEs to be used.

The rest of this chapter is structured as follows. The formalisation and algorithmic definition of the proposed MDPOHMT is presented in Section 4.2. A novel design tool for specifying the

degree of relaxation to be applied in the presence of varying noise is presented in Section 4.3, the applicability of the MDPOHMT on N-channel imagery is discussed briefly in Section 4.4 and, finally, the results of using the MDPOHMT are presented, as well as comparisons with other, similar, colour HMTs in Section 4.5.

4.2 Extension of the Morphological Hit-or-Miss Transform

Similar to the methods described in [40, 67], the approach proposed in this chapter, the MDPOHMT, uses a pair of N-dimensional, colour or spectral, SEs. Other approaches to extend morphology to multivariate images aim to use a traditional two dimensional support [69, 74, 267, 268]. However, as the ultimate goal of the HMT is to detect objects of interest based on their size and shape, providing an additional parameter, in the form of an object's spectral information where appropriate, is desirable in many applications and removes some of the colour ambiguity issues faced by other methods. A single pixel in an N-channel multivariate image can be considered as a point in a N-dimensional space, $\mathbb{Z}^N$. Each pixel within the support of a SE of the same dimensionality, which has been translated to some point in the image, can be represented by a point in the same space, some distance, d, from its corresponding image pixel.

The fundamental theory behind object detection using the morphological HMT, is that, for any pixel, x, to be considered in the set of detections, every pixel in the region of an image bounded by the support of $\boldsymbol{S}_{FG}$, when translated to x, matches $\boldsymbol{S}_{FG}$. Whilst simultaneously, every pixel in the region of an image bounded by the support of $\boldsymbol{S}_{BG}$, when translated to x, is "missed" by $\boldsymbol{S}_{BG}$.

In binary and greyscale morphology, this notion of matching is formalised using set theory, as discussed in Section 3.2.5. Whilst further extension to colour and multi-channel images cannot easily be characterised using set theory alone, the concept that $\boldsymbol{S}_{FG}$ must match the image better than $\boldsymbol{S}_{BG}$ can be expressed using the notion of distance within a multi-dimensional space. Specifically, given an image, $\boldsymbol{X} \in \mathbb{Z}^N$ and a composite SE, $\boldsymbol{S} \in \mathbb{Z}^N$, the distance, d, between each image pixel and its corresponding pixel in the SE, when $\boldsymbol{S}$ is translated to some point $x \in \boldsymbol{X}$, can be used as a measure of how well each pixel matches. The degree to which each element in both $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ match the image can easily be determined.

The result of a multi-channel HMT in this space can be expressed as the locations, where $\boldsymbol{S}_{FG}$ matches the image, *i.e.,* exhibits low values of d between each pixel in $\boldsymbol{X}$ within the support

of $\boldsymbol{S}_{FG}$ when translated to some point $x \in \boldsymbol{X}$. Simultaneously, $\boldsymbol{S}_{BG}$ should "fit" around the image, characterised as being further from matching the image within the space. This can be similarly expressed as higher values of d between each pixel in $\boldsymbol{S}_{BG}$, when translated to some pixel, x, and the corresponding image pixel.

This means that, without any relaxation to account for noise or other occlusion effects, the result of the MDPOHMT is the locus of all points in the image where the maximum (*supremum*) of the set of foreground distances, d_{FG}, is less than the minimum (*infimum*) of the set of background distances, d_{BG}. This idea is shown in Figure 4.1 with seven scenarios where the MDPOHMT seeks to detect a 3×3 red square on some image plane containing a blue and red square, each of size 3×3.

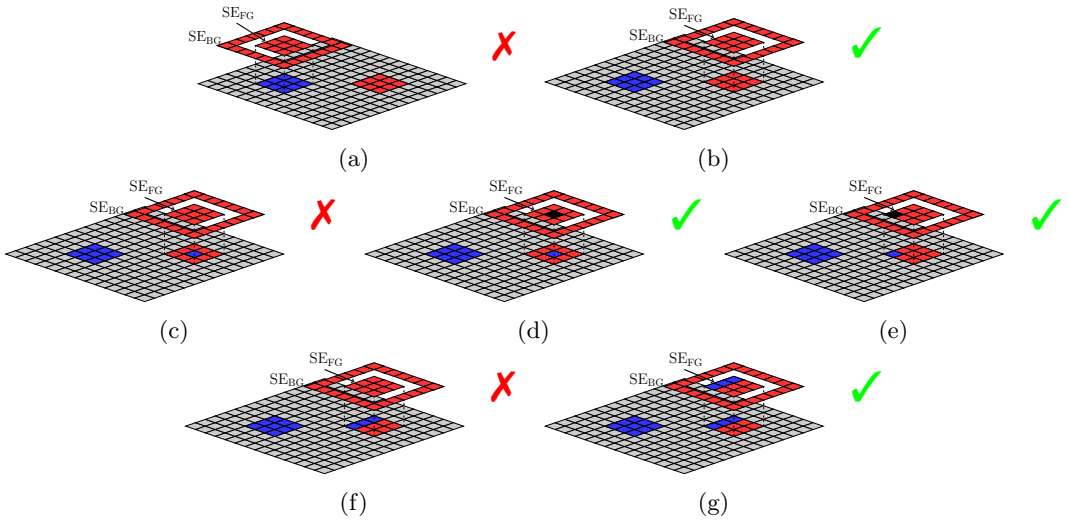


Figure 4.1: An example of the use cases of the MDPOHMT in pictorial form.

In each of the scenarios shown in Figure 4.1, the composite SE $\boldsymbol{S} = [\boldsymbol{S}_{FG}, \boldsymbol{S}_{BG}]$, consisting of a 3×3 square red $\boldsymbol{S}_{FG}$ and a 7×7 red hollow square $\boldsymbol{S}_{BG}$ with a thickness of 1 pixel, is centred above some object of interest, in this case red and blue 3×3 squares. At each location in the image, the question of how well the foreground and background SEs match the image is asked by calculating a set of dissimilarities between corresponding pixels in the image and SE. The MDPOHMT is then simply the set of pixels in which the foreground SE matches the image better than the background SE at some PO value.

In each example shown in Figure 4.1, the minimum background dissimilarity, d_{BG}, is identical. The maximum foreground dissimilarity, d_{FG}, varies based on the colours within the support of $\boldsymbol{S}_{FG}$. However, it can be noted that $d_r < d_{BG} < d_b$, where d_b and d_r are the dissimilarities between the red pixels in $\boldsymbol{S}_{FG}$ and the blue and red image pixels respectively. In

Figure 4.1a, $\boldsymbol{S}$ is translated such that $\boldsymbol{S}_{FG}$ is centred on the blue object and $\boldsymbol{S}_{BG}$ fits around it. The dissimilarity between the image and $\boldsymbol{S}_{FG}$ is larger than the dissimilarity between $\boldsymbol{S}_{BG}$ and the image which results in no detection. Conversely, in Figure 4.1b, there is no dissimilarity between the image and $\boldsymbol{S}_{FG}$. As a result, $d_{FG} < d_{BG}$ and the object is detected. In Figure 4.1c a single pixel of noise, in the form of a flipped blue pixel, will cause the MDPOHMT to "miss" the desired object, as d_{FG} would be identical to that of Figure 4.1a. However, by incorporating some form of PO relaxation and instead taking the second largest dissimilarity as the value of d_{FG}, the object can once again be detected. This is shown in Figure 4.1d, where the allowance for a "don't care" pixel in $\boldsymbol{S}_{FG}$ is given in black. This relaxation is unaffected by the location of the noise as shown in Figure 4.1e. To achieve this, SEs should be designed such that they have a general size, shape, and colour, along with an acceptable provision for noise in the form of a PO value. In Figure 4.1f, the object to be detected has changed and therefore it cannot be detected using $\boldsymbol{S}$. Rather than excessively relaxing the transform, $\boldsymbol{S}_{FG}$ can be changed and the points in multidimensional space can be updated, matching the object once again, and it can be detected in Figure 4.1g highlighting that non-flat colour SEs can be used as $\boldsymbol{S}_{FG}$ does not need to be a single homogeneous colour, and in some cases this is advantageous for detecting more complex objects.

The proposed MDPOHMT is the combination of a pair of colour SEs along with a pair of appropriate rank-order filters which reduce its susceptibility to misidentification due to noise, occlusion and other common issues faced in MM.

4.2.1 Development of a Novel Reduced Ordering-based Multivariate Hit-or-Miss Transform

The algorithm described pictorially in Figure 4.1 is formalised here and considers the set of images $\mathcal{I}$ from $\mathbb{Z}^n$, the discrete domain of dimension $n > 0$ into the interval $[0,1]^n$: *i.e.,* $\mathcal{I} = \mathbb{Z}^n \rightarrow [0,1]^n$. Given an image $\boldsymbol{I} \in \mathcal{I}$, the support of $\boldsymbol{I}$, denoted by $\text{supp}(\boldsymbol{I})$, is the set of points where $\boldsymbol{I}$ is non-zero: *i.e.,* $\text{supp}(\boldsymbol{I}) = \{x \in \mathbb{Z}^n, \boldsymbol{I}(x) \neq 0\}$. In the following, it is always assumed that the considered images have a finite support. In order to perform the MDPOHMT in this space, measurements between $\boldsymbol{I}$ and a probing composite SE, $\boldsymbol{S}$, must be defined.

Let $\hat{d}$ be a normalised positive dissimilarity measure on $[0,1]^n$. The Euclidean distance function provides a general measure for dissimilarity between two points. This can be normalised, *i.e.,* bounded in the range 0–1, by dividing the dissimilarity values, d, by its maximum possible value, effectively turning the maximum dissimilarity into a unit vector. In the case of the Euclidean distance, this is the square root of the number of channels present

in the image or dimensions in $\mathbb{Z}^n$, n. The normalised, Euclidean distance-based, dissimilarity, $\hat{d}$, between two points $a, b \in \mathbb{Z}^n$ is calculated as

$$\hat{d}(a,b) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(a_i - b_i)^2}. \tag{4.1}$$

It should be noted that the use of the Euclidean distance is merely an example of a distance metric, there are multiple other distance measures which are sometimes more suited to particular applications as discussed in Chapter 2.

Given two images $\boldsymbol{A}$ and $\boldsymbol{B}$ in $\mathcal{I}$ and a point, x, in $\mathbb{Z}^n$ with some method for measuring normalised distance between each corresponding pair of pixels in $\boldsymbol{A}$ and $\boldsymbol{B}$, the indexed family of normalised dissimilarities, DS, between $\boldsymbol{A}$ and $\boldsymbol{B}$ at point x can be defined as:

$$DS(\boldsymbol{A}, \boldsymbol{B}, x) = \left\{\hat{d}(\boldsymbol{A}(x+b), \boldsymbol{B}(b))\right\}_{b \in \text{supp}(\boldsymbol{B})} \tag{4.2}$$

Replacing $\boldsymbol{A}$ and $\boldsymbol{B}$ with an image $\boldsymbol{I}$ and either $\boldsymbol{S}_{BG}$ or $\boldsymbol{S}_{FG}$ results in the set of background or foreground dissimilarities at a point, x, respectively. By ordering these sets, it is possible to obtain the rank k dissimilarity in an identical way to that of the greyscale intensities in Equation (3.20) based on some SE, $\boldsymbol{SE}$, and a desired rank, k.

$$[\xi_{\boldsymbol{SE},k}(\boldsymbol{I})](x) = k^{\text{th}} \underset{s \in \boldsymbol{SE}}{\text{order statistic}} \{DS(\boldsymbol{I}, \boldsymbol{SE}, x)\} \tag{4.3}$$

In some cases, it may be desirable, or advantageous, to alter the constraints on either SE individually, to have increased control over the levels of noise or occlusion acceptable in the transform, whilst still detecting objects of interest. This can be achieved by altering the PO values for the foreground and background independently. As a result, the rank of each filter k_{FG} and k_{BG}, is based on the desired PO value for foreground and background such that $50\% < p_{FG},\ p_{BG} \leq 100\%$ and can be calculated as

$$k_{FG} = \text{Card}(\boldsymbol{S}_{FG}) - \left\lfloor \left(1 - \frac{p_{FG}}{100}\right) \text{Card}(\boldsymbol{S}_{FG}) \right\rfloor \tag{4.4}$$

$$k_{BG} = 1 + \left\lfloor \left(1 - \frac{p_{BG}}{100})\right) \text{Card}(\boldsymbol{S}_{BG}) \right\rfloor \tag{4.5}$$

The range of PO values must lie between 50% and 100% as 50% occupancy would simply be a median filter and setting the occupancy to any value $< 50\%$ would result in an erosion tending towards a dilation and vice versa. Setting the PO value to 100% would result in

$k_{FG} = \text{Card}(\boldsymbol{S}_{FG})$, *i.e.,* a dilation or maximum filter, and $k_{BG} = 1$, *i.e.,* an erosion or minimum filter. Using Equation (4.3), both a relaxed dilation of the foreground dissimilarities and a relaxed erosion of the background dissimilarities can be defined. By combining these with a variable foreground similarity threshold, $T \in [0, 1]$, a new multivariate HMT-like operator, the MDPOHMT, is defined in Equation (4.6),

$$[\text{MDPOHMT}_{\boldsymbol{S}}(\boldsymbol{I})](x) = \begin{cases} 1, & \text{if } [\xi_{\boldsymbol{S}_{FG},k_{FG}}(\boldsymbol{I})](x) < [\xi_{\boldsymbol{S}_{BG},k_{BG}}(\boldsymbol{I})](x) \\ & \text{and } [\xi_{\boldsymbol{S}_{FG},k_{FG}}(\boldsymbol{I})](x) \leq T \\ 0, & \text{otherwise,} \end{cases} \tag{4.6}$$

where k_{FG} and k_{BG} are the appropriate ranks based on some desired PO values calculated using Equations (4.4) and (4.5) respectively. In Equation (4.6), whilst similar to Equation (3.22), the polarity of the operations on the foreground and background are switched, with the relaxed erosion of foreground intensities becoming a relaxed dilation of foreground dissimilarities and, similarly, the relaxed dilation of background intensities becomes a relaxed erosion of background dissimilarities.

If $p_{FG} = p_{BG} = 100\%$, Equation (4.6) reduces to determining if $\max(DS(\boldsymbol{I}, \boldsymbol{S}_{FG}, x)) < \min(DS(\boldsymbol{I}, \boldsymbol{S}_{BG}, x))$, *i.e.,* the largest dissimilarity between the image and $\boldsymbol{S}_{FG}$ is compared to that of the smallest dissimilarity between the image and $\boldsymbol{S}_{BG}$. If the image is close to $\boldsymbol{S}_{FG}$ and far from $\boldsymbol{S}_{BG}$ when translated to an object of interest, there is a larger disparity between d_{FG} and d_{BG} which indicates the SE is suited to detecting this object. Additionally, noise robustness is improved by providing a buffer to account for any potential degradation of the distance between d_{FG} and d_{BG}, *e.g.,* with the application of mean-centred Gaussian noise.

The operation of the MDPOHMT is displayed in Figure 4.2. Figure 4.2a shows an example where a red object is being probed with a red composite SE, $\boldsymbol{S}$. Plotting each point within the supports of both $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$, as well as the values in the SEs themselves, results in Figure 4.2b, where d_{FG} and d_{BG} can be visualised in RGB space. As there is no dissimilarity between the image and $\boldsymbol{S}_{FG}$ when centred at the point under consideration, d_{FG} is zero. The minimum background dissimilarity, d_{BG}, however, is non-zero and is simply the Euclidean distance between the red pixels in $\boldsymbol{S}_{BG}$ and the orange image background. In Figure 4.2c, $\boldsymbol{S}$ is centred on a blue object instead. As a result, in Figure 4.2d, d_{FG} is no longer zero and is instead the Euclidean distance between red and blue in the RGB space, whilst d_{BG} is unchanged. As $d_{FG} > d_{BG}$, this object is not detected. Figure 4.2e shows $\boldsymbol{S}$ centred on a red object punctured by noise in the form of a single blue pixel. While most of the pixels within the support of $\boldsymbol{S}_{FG}$ match the object, the presence of this single blue pixel is enough to cause the maximum

foreground dissimilarity, d_{FG}, to be identical to that of the blue object in Figure 4.2c as shown in Figure 4.2f. This highlights the need for PO-based relaxation, as a single pixel of noise is sufficient to cause the MDPOHMT to fail to detect objects of interest. Finally, Figure 4.2g shows an example where the object from Figure 4.2a has been corrupted with Gaussian noise with mean $\mu = 0$, variance $\sigma^2 = 0.05$, and inter-channel correlation $\rho = 0$. This results in no pixel in either $\boldsymbol{S}_{FG}$ or $\boldsymbol{S}_{BG}$ having an exact fit with the image as shown in Figure 4.2h. The noise is such that $d_{FG} > d_{BG}$ by a small margin meaning that, despite the apparent compatibility of the object and the SE, it cannot be detected without some form of relaxation.

The sets of dissimilarities for each of the scenarios in Figure 4.2 are shown in Figure 4.3. From top to bottom, Figure 4.3a shows the colour values within the support of $\boldsymbol{S}_{FG}$ when centred on the object from Figure 4.2a, the Euclidean distance-based normalised dissimilarity between the image and $\boldsymbol{S}_{FG}$, the ordered dissimilarity values, and the ordered image colour values. As the image and $\boldsymbol{S}_{FG}$ match exactly, the dissimilarity at each location is zero. Similarly, Figures 4.3b and 4.3c show the same values when $\boldsymbol{S}_{FG}$ is centred at the objects shown in Figures 4.2c and 4.2e respectively. In each case, the foreground dissimilarity, d_{FG}, has a value of 0.82 as the maximum dissimilarity between the image and probing $\boldsymbol{S}_{FG}$ is between red and blue pixels. Figure 4.3d shows the image and SE colours within the support of $\boldsymbol{S}_{BG}$ as well as their dissimilarities for each of the scenarios shown in Figures 4.2a, 4.2c and 4.2e. As the maximum dissimilarity between the $\boldsymbol{S}_{FG}$ and image, from Figure 4.3a, is less than the minimum dissimilarity between $\boldsymbol{S}_{BG}$ and the image in Figure 4.3d, the object in Figure 4.2a is detected. However, the opposite is true when comparing Figures 4.3b and 4.3c with Figure 4.3d and therefore these objects are not detected. Figures 4.3e and 4.3f show the SE values and image values within the support of $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ respectively when centred on the noisy example in Figures 4.2g and 4.2h. Due to the additive noise, $d_{BG} = 0.21$, while, $d_{FG} = 0.22$, resulting in no match, highlighting the need for relaxation in noisy or non-ideal objects.

Relaxing the required fit of the MDPOHMT by reducing $p_{_{FG}}$ or $p_{_{BG}}$ leads to an increased True Positive Rate (TPR) but can also result in an increased False Positive Rate (FPR) as the relaxation can allow for both desired and undesired objects to be detected more readily. As well as the rank parameter, the threshold, T, can be adjusted to specify a maximum dissimilarity between the image and the corresponding SE that can register a hit, this can account for variability of colour in objects. For example, if the target of interest is red, the threshold can be set to ensure that a probing red $\boldsymbol{S}_{FG}$ only "fits" with colours similar enough to a potential match and rules out other unwanted colours that still match better than their background.

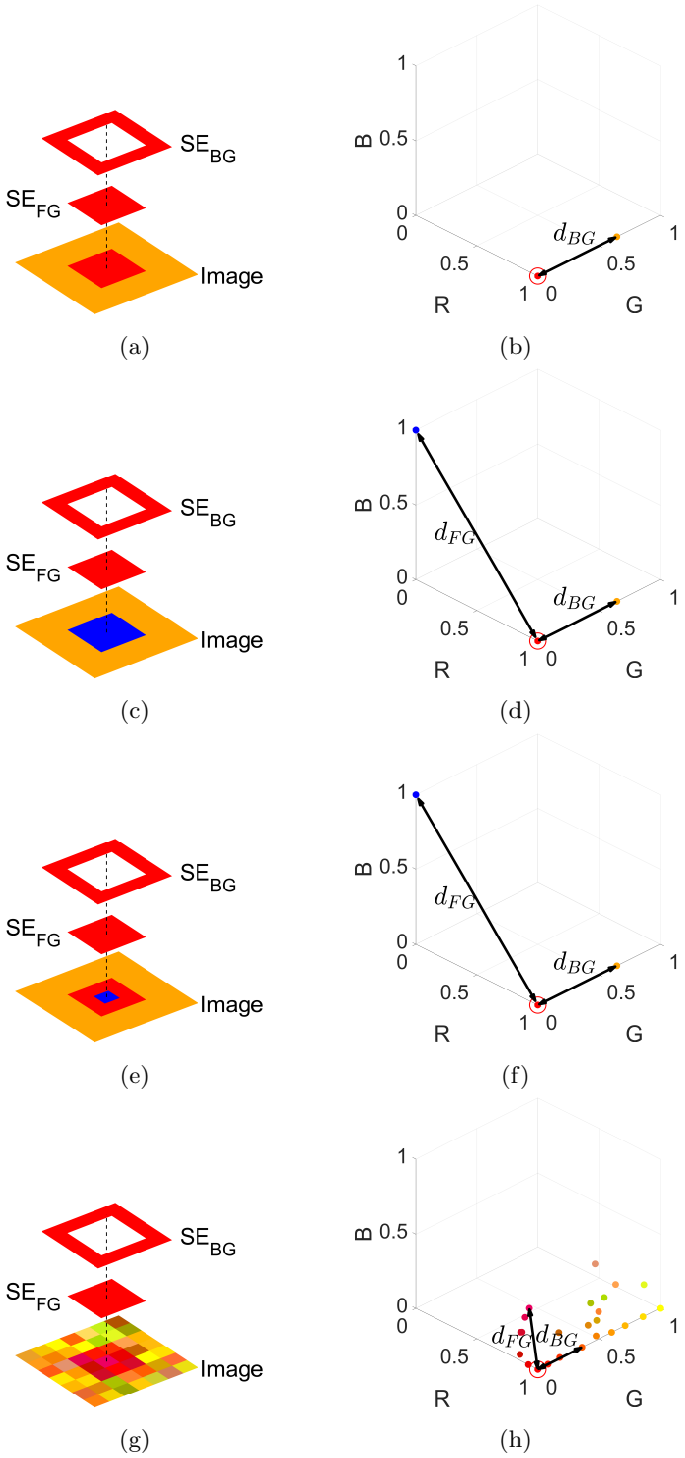


Figure 4.2: Example foreground and background dissimilarities in the RGB colour space **a)** True positive example. **b)** Resultant d_{FG} and d_{BG} of Figure 4.2a. **c)** True negative example. **d)** Resultant d_{FG} and d_{BG} of Figure 4.2c. **e)** False negative example (flipped). **f)** Resultant d_{FG} and d_{BG} of Figure 4.2e. **g)** False negative example (noise). **h)** Resultant d_{FG} and d_{BG} of Figure 4.2g.

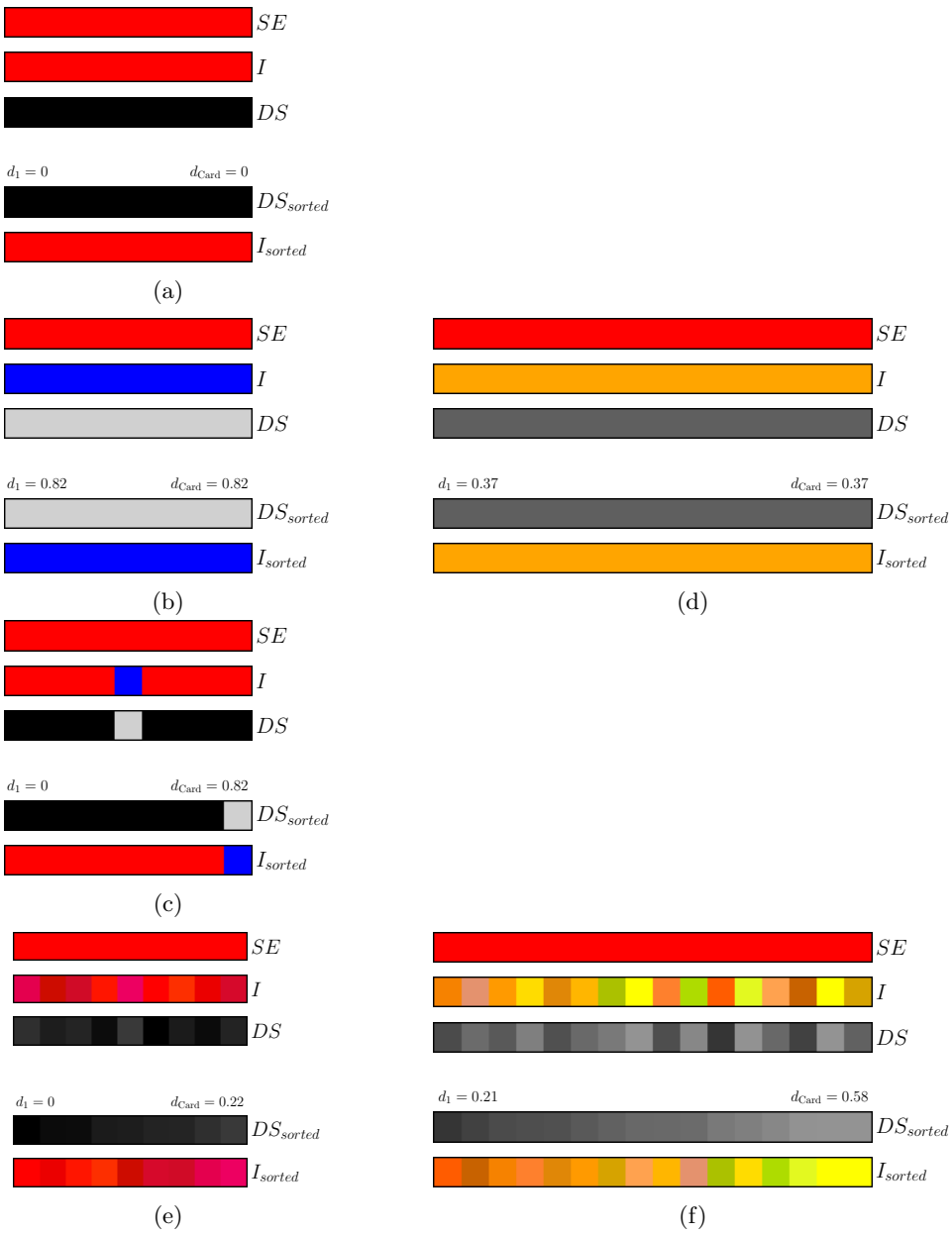


Figure 4.3: Ordered dissimilarities from Figure 4.2 **a)** Foreground dissimilarity from Figure 4.2a. **b)** Foreground dissimilarity from Figure 4.2c. **c)** Foreground dissimilarity from Figure 4.2e. **d)** Background dissimilarities from Figures 4.2a, 4.2c and 4.2e. **e)** Foreground dissimilarity from Figure 4.2g. **f)** Background dissimilarity from Figure 4.2g.

4.3 Optimal Parameter Setting for Noise Robustness in a Multivariate Hit-or-Miss Transform

It has been shown previously that rank order filters can be used in order to add robustness to both binary and greyscale HMTs in the presence of noise. The PO plot method defined in [50] can be used to estimate the most appropriate rank from example objects and indeed can be extended for use with the MDPOHMT. When centred at a pixel in a greyscale image, the intensities of the image pixels are compared with their corresponding SE values. The occupancy can then be calculated for foreground and background by determining the percentage of pixels that lie below some grey-level, t, in $\boldsymbol{S}_{FG}$ and the percentage that are above t in $\boldsymbol{S}_{BG}$. Simply put, this represents how much of $\boldsymbol{S}_{FG}$ fits within the foreground and how much of $\boldsymbol{S}_{BG}$ fits around it at any value of t. Obtaining these occupancy values for each value of t produces two occupancy vectors containing the foreground and background occupancies respectively. When plotted against one another, these two vectors produce a PO plot at the particular pixel of interest.

When applied to multivariate images, these estimated occupancies need to relate to colour and multivariate dissimilarity as opposed to greyscale intensity. This, along with the lack of grey-level analogue in a colour image, means a slightly different approach is required in order to extend the notion of PO plots for use with multivariate images. This extension of the PO plot rank parameter setting method to the multivariate domain is one of the main contributions of this work and provides a simple and effective approach to improve the recognition rate of objects and patterns in the presence of noise.

In order to obtain the occupancy vectors using multivariate images and SEs, the Probability Density Function (PDF) of the ordered normalised dissimilarity for each SE is first estimated. This is achieved by binning each of the sets of foreground and background dissimilarities to create the histograms H_{FG} and H_{BG} respectively. The PDFs of dissimilarities between the image and each SE are estimated by normalising each histogram using Equation (4.7)

$$\mathrm{PDF}(DS) \approx \hat{H} = \frac{1}{\mathrm{Card}(\boldsymbol{SE})} \cdot H, \tag{4.7}$$

where H is the binned set of dissimilarities, DS, between an image and some SE, $\boldsymbol{SE}$. The occupancy vectors PO_{FG} and PO_{BG} are then calculated from the Cumulative Distribution Function (CDF) of these estimated PDFs using Equations (4.8) and (4.9):

$$PO_{FG}(n) = \text{CDF}(FG) = \sum_{m=1}^{n} \hat{H}_{FG}(m), \tag{4.8}$$

$$PO_{BG}(n) = 1 - \text{CDF}(BG) = 1 - \sum_{m=1}^{n} \hat{H}_{BG}(m). \tag{4.9}$$

Here, the occupancy vectors indicate whether $\boldsymbol{S}_{FG}$ or $\boldsymbol{S}_{BG}$ is more similar to the image at any given rank. Following the notion that $\boldsymbol{S}_{FG}$ must fit the image better than $\boldsymbol{S}_{BG}$, a critical point corresponding to the minimum occupancy required to detect an object can be found as $P = \max\left[\min\{PO_{FG}, PO_{BG}\}\right]$. This critical point can be visualised by plotting both PO_{FG} and PO_{BG} together and finding their point of intersection. Plotting PO_{FG} against PO_{BG} producing a PO plot also highlights this critical point as the location of the graphs intersection with the straight line $PO_{FG} = PO_{BG}$.

An example of this can be seen in Figure 4.4 for clean (Figure 4.4a), flipped (Figure 4.4d), and noisy (Figure 4.4g) test images respectively. The flipped image was created by altering the clean image using the existing colour palette, ensuring that some objects were punctured. The noisy image was created by then adding Gaussian noise with $\mu = 0$, $\sigma^2 = 0.05$, and $\rho = 0$ to the flipped image.

The occupancy graphs for each of the test images are shown in Figures 4.4b, 4.4e and 4.4h. These occupancy plots indicate the proportion of pixels within the support of $\boldsymbol{S}_{FG}$ that are less than the corresponding dissimilarity from the image, while simultaneously displaying the proportion of pixels in the support of $\boldsymbol{S}_{BG}$ that are greater than that same dissimilarity from the image. By finding the maximum of the minimum of all points in PO_{FG} and PO_{BG}, or the point of intersection between the two, (indicated by the horizontal black lines in Figures 4.4b, 4.4e and 4.4h) an estimate of PO can be found. In both the clean and noisy cases, the object of interest is the red square in the upper left of each image, whereas, in the flipped image, it is the hollow red square in the centre of the image. Each object of interest in the images for which the PO plots are generated is highlighted with a black ring in Figures 4.4a, 4.4d and 4.4g.

In the case of the clean object (Figure 4.4a) the required PO value for a successful detection using the MDPOHMT was, as expected, 100%. For the object in the flipped image (Figure 4.4d), 1/9th of the area of the object had been edited from red to orange. This resulted in a PO value of 88.89%, (*i.e.,* $\sim$ 8/9ths) in order to detect it successfully. For the object corrupted by Gaussian noise (Figure 4.4g) the PO value required to successfully detect it using the MDPOHMT fell to 87%.

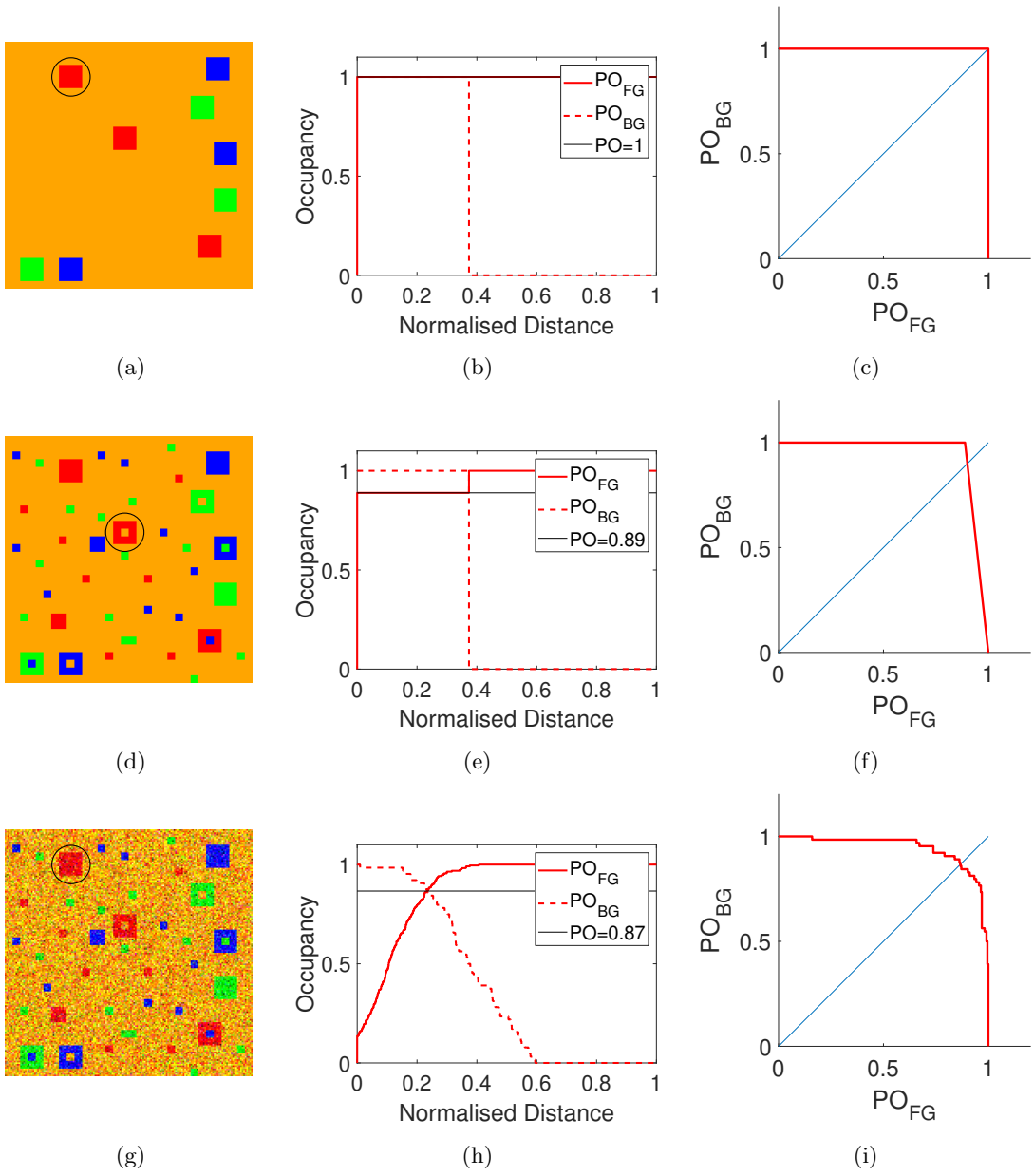


Figure 4.4: PO plots and setting of the appropriate PO and rank parameters **a)** Clean 160×160 test image. **b)** Occupancy of $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ translated to the point in Figure 4.4a. **c)** PO plot of the $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ translated to the point in Figure 4.4a. **d)** Flipped 160×160 test image. **e)** Occupancy of $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ translated to the point in Figure 4.4d. **f)** PO plot of the $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ translated to the point in Figure 4.4d.**g)** Noisy 160×160 test image. **h)** Occupancy of $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ translated to the point in Figure 4.4g. **i)** PO plot of the $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ translated to the point in Figure 4.4g.

Producing PO plots for each of these images at the marked locations results in the curves shown in Figures 4.4c, 4.4f and 4.4g By analysing where each curve crosses the diagonal the optimal PO value can be estimated. The PO plots produce the same values of PO as the graphs from Figures 4.4b, 4.4e and 4.4h. Once the PO estimation has been made using either method, the rank parameters, k_{FG} and k_{BG}, can be found for each SE using Equations (4.4) and (4.5) respectively.

4.4 Applying a Multivariate Hit-or-Miss Transform to Hyperspectral Target Detection

Hyperspectral imaging is becoming increasingly prevalent in academic and industrial applications and the MDPOHMT operator was designed with this in mind. The MDPOHMT has no inherent restriction on dimensionality, as only the dissimilarity metric used dictates the input dimensionality of the data. As a result, it can be generalised for 3D images with N channels such as volumetric, multispectral, and hyperspectral images with the design of an appropriate SE and choice of distance measure. In the case of hyperspectral TD, both the spectrum and desired shape(s) of target instances are known and therefore SEs can be designed accordingly. Spectral detection algorithms such as the SAM and SID algorithms can replace the Euclidean distance as the dissimilarity function of the MDPOHMT. Alternatively, DR may be carried out on both the SE and image with the detection performed in the reduced domain similar to the techniques described in Chapter 5.

4.5 Results

In this section, the MDPOHMT is applied to various synthetic and natural multivariate images. The MDPOHMT is first compared to an existing greyscale HMT in order to highlight the need for a generalisation to multivariate imagery to better distinguish between objects of interest. The benefits of using PO as a method for noise robustness when combined with the MDPOHMT is highlighted by comparing its performance to the colour SHMT. Examples of object detection in high-resolution aerial remote sensing images are shown, again comparing the MDPOHMT to the SHMT considering objects of varying size, shape, colour, and texture. Finally, examples of extending the MDPOHMT to higher dimensionality multivariate images are shown by performing object detection in hyperspectral imagery, using both synthetic data and natural images.

4.5.1 Image Collection

A combination of synthetic data and natural images are used in order to test the performance of the MDPOHMT and compare it against other, similar, HMTs. Initially, synthetic images were created to easily analyse and verify the behaviour of the MDPOHMT. Following this, multiple natural images were used to validate the performance of the MDPOHMT when not applied to ideal imagery.

In order to test the performance of the MDPOHMT in detecting objects in numerous, high resolution, images the DOTA dataset [269] was used. The full DOTA dataset contains thousands of images from Google Earth with 15 object classes to be detected. For the purposes of testing the MDPOHMT, a subset of these images which contain swimming pools are used as they have varying size, shape, orientation, and colour information. Additionally, each instance has a varying amount of pixels on target, providing a complex set of target objects. The dataset contains a ground truth for its training and validation subsets, ground truths for further images were created in order to increase the volume of data available to test on. In total, 7 high resolution colour images were selected for testing with a total of 1257 instances of the swimming pool object included.

In order to test the extension of the MDPOHMT beyond RGB imagery for application on MSI and HSI data, synthetic images were created using methods from the Hyperspectral Imagery Synthesis toolbox [270] using spectra from the USGS Spectral Library [271]. Finally, the MDPOHMT is tested on a set of images supplied by BAE systems containing multiple targets captured from an aerial platform.

4.5.2 Synthetic Colour Image Tests

Figure 4.5 shows the clean, flipped, and noisy test images from Figures 4.4a, 4.4d and 4.4g, as well as PO plots when an appropriate SE, Figures 4.5m to 4.5o, is translated to each of the red, green, and blue objects in each image respectively.

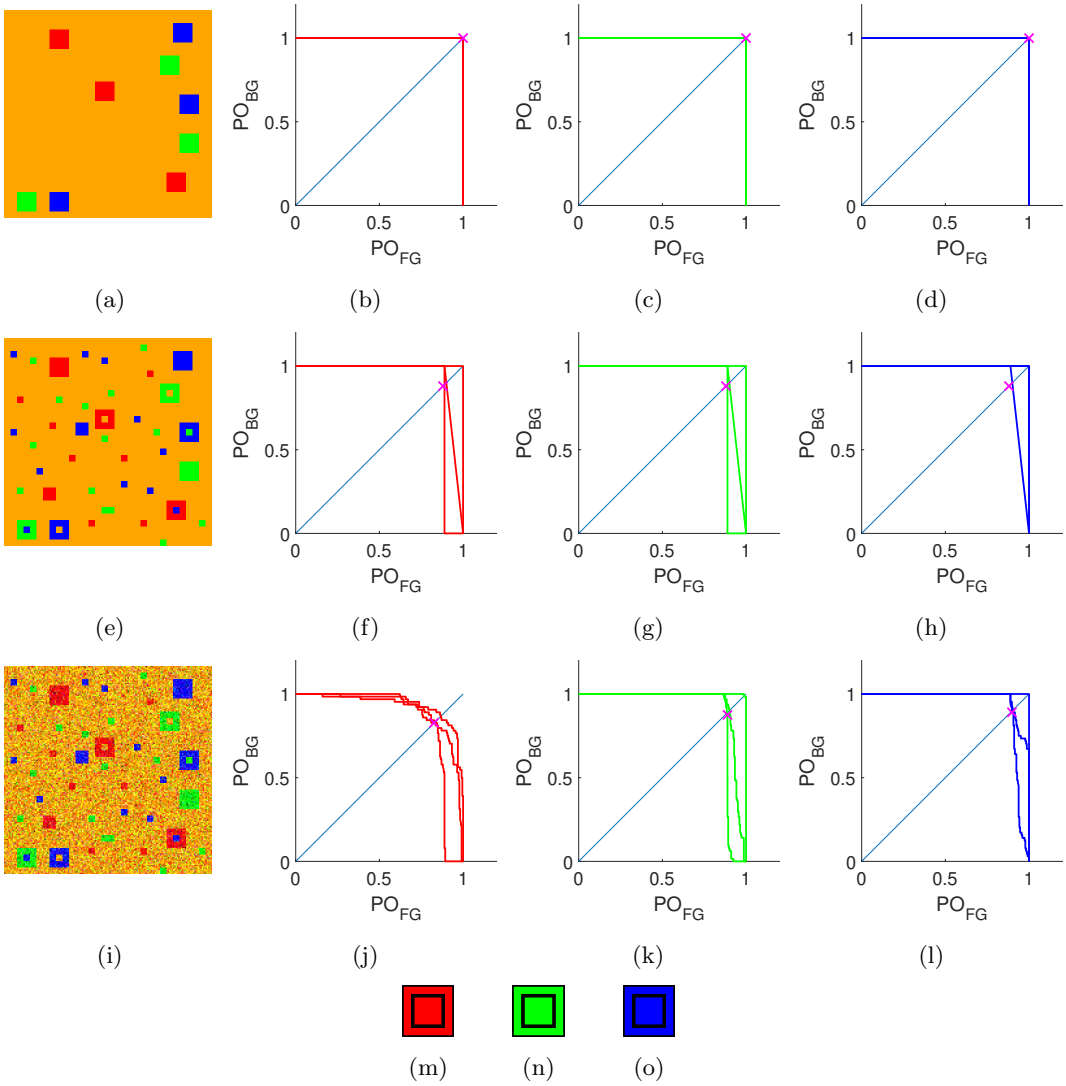


Figure 4.5: Percentage occupancy plots of coloured square in test images. **a)** Clean 160×160 test image. **b)** PO plots for the red objects, $p = 100\%$. **c)** PO plots for the green objects, $p = 100\%$. **d)** PO plots for the blue objects, $p = 100\%$. **e)** Flipped 160×160 test image. **f)** PO plots for the red objects, $p = 87\%$. **g)** PO plots for the green objects, $p = 88\%$. **h)** PO plots for the blue objects, $p = 88\%$. **i)** Noisy 160×160 test image. **j)** PO plots for the red objects, $p = 82\%$. **k)** PO plots for the green objects, $p = 88\%$. **l)** PO plots for the blue objects, $p = 89\%$. **m)** Red SE. **n)** Green SE. **o)** Blue SE.

In the case of the clean image (Figure 4.5a) all nine objects exhibit a perfect fit with their respective SEs, as shown in Figures 4.5b to 4.5d. As each of the SEs match the objects, this

produces square PO plots which, as the plots intersect with the top right corner *i.e.,* $(1, 1)$ or $PO_{FG} = PO_{BG} = 100\%$, indicates a perfect fit. Similar to the example shown in Figure 4.5i, each of the objects in the flipped image (Figure 4.5e), with the exception of the three unchanged objects, have their centres flipped, corresponding to 1/9th of their area. As a result the PO value required to detect all of the objects in the flipped image is 88%, as shown in Figures 4.5f to 4.5h.

The noisy image is identical to the one shown in Figure 4.4g. The required PO values at the centre of each object in the image were calculated using the PO plots shown in Figures 4.5j to 4.5l. In the noisy case, the PO plot is no longer square, indicating a non-exact fit. The required PO, and resultant rank estimation, can be found through investigation of these plots. Selecting the minimum occupancy from the set of points queried, and obtaining appropriate rank values for $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$, is sufficient to ensure these objects are detected using the MDPOHMT.

Applying the MDPOHMT to the test image shown in Figure 4.5a with 100% occupancy results in an ideal fitting, as expected through analysis of the PO plots and verified by the MDPOHMT result shown in Figures 4.6a, 4.6d and 4.6g for the red, green, and blue objects respectively. The foreground and background distance images, used in calculating the MDPOHMT, are also shown in Figures 4.6c and 4.6d, 4.6e and 4.6f, and 4.6h and 4.6i for the red, green and blue SEs respectively.

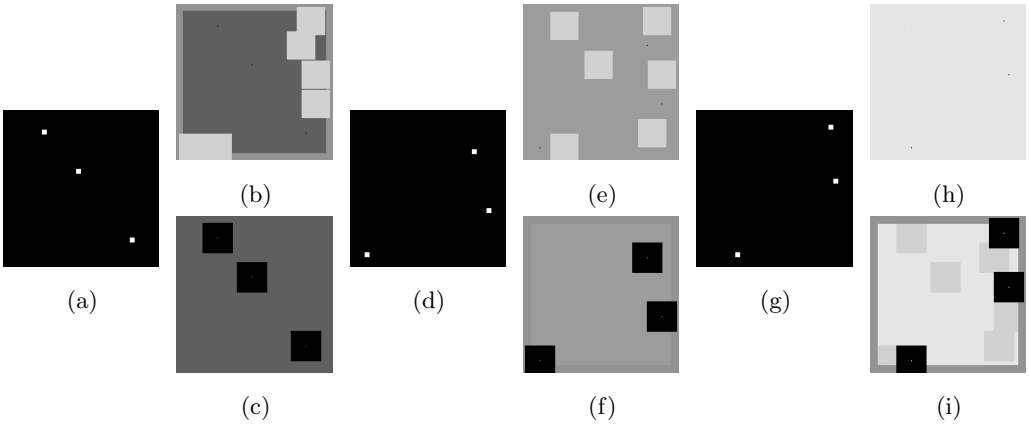


Figure 4.6: Results of the MDPOHMT on a clean synthetic test image. **a-c)** MDPOHMT, FG dissimilarity, and BG dissimilarity of Figure 4.5a using Figure 4.5m with $p = 100\%$. **d-f)** MDPOHMT, FG dissimilarity, and BG dissimilarity of Figure 4.5a using Figure 4.5n with $p = 100\%$. **g-i)** MDPOHMT, FG dissimilarity, and BG dissimilarity of Figure 4.5a using Figure 4.5o with $p = 100\%$.

In each of the foreground distance images (Figures 4.6b, 4.6e and 4.6h) there are three pixels of zero dissimilarity where each of the red, green, and blue foreground SEs completely intersect

with the corresponding objects on the image plane. At the same positions in the background dissimilarity images (Figures 4.6c, 4.6f and 4.6i), there are single pixels of high dissimilarity, where the background SEs fit around each of the objects of interest. These pixels are surrounded by areas of zero dissimilarity as in each case, $\boldsymbol{S}_{BG}$ intersects with the object in the foreground. The resultant MDPOHMT has been dilated using a 5×5 square SE for visualisation purposes, but represents a single pixel at the centre of each object being detected.

Applying the MDPOHMT with the same parameters to the noise corrupted image (Figure 4.5i) results in only one of the blue objects being detected, as visible in Figures 4.7a, 4.7d and 4.7g. This is because the each foreground SE no longer fits the image as it is being "punctured" by noise, increasing d_{FG}. Similarly, noise in the background can lead to a decrease in d_{BG}.

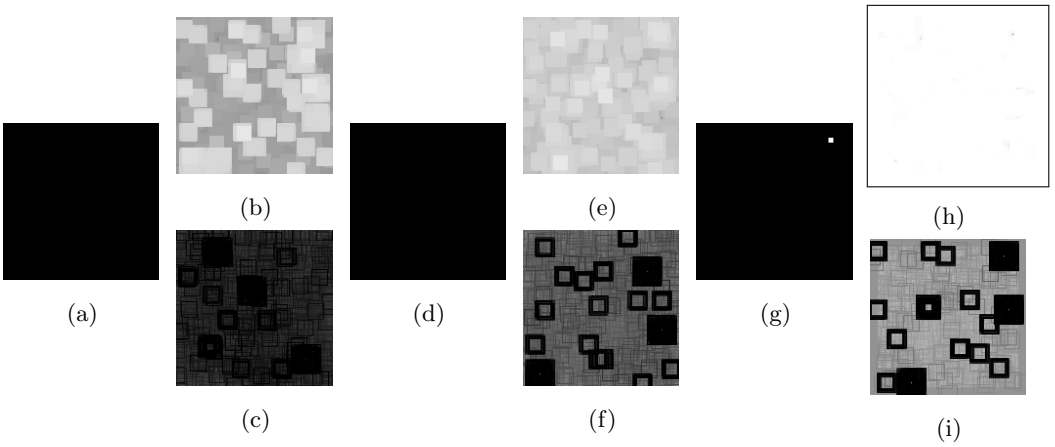


Figure 4.7: Results of the MDPOHMT on a noisy synthetic test image. **a-c)** MDPOHMT, FG dissimilarity, and BG dissimilarity of Figure 4.5d using Figure 4.5m with $p = 100\%$. **d-f)** MDPOHMT, FG dissimilarity, and BG dissimilarity of Figure 4.5d using Figure 4.5n with $p = 100\%$. **g-i)** MDPOHMT, FG dissimilarity, and BG dissimilarity of Figure 4.5d using Figure 4.5o with $p = 100\%$.

Using the PO plots from Figures 4.5j to 4.5l, the required PO necessary to detect each of these objects is found to be 82%, 88%, and 89% for the red, green, and blue objects respectively. These PO values inform the rank parameter required to account for noise in $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ in each object. Applying the MDPOHMT with these relaxed parameters, the objects are detected correctly as shown in Figures 4.8a, 4.8d and 4.8g. The foreground (Figures 4.8b, 4.8e and 4.8h) and background (Figures 4.8c, 4.8f and 4.8i) dissimilarity images are also shown where it is possible to observe this relaxation has the effect of decreasing d_{FG} at each pixel, or darkening the image, and increasing d_{BG} at every pixel, or lightening the image, compared to those in Figure 4.7.

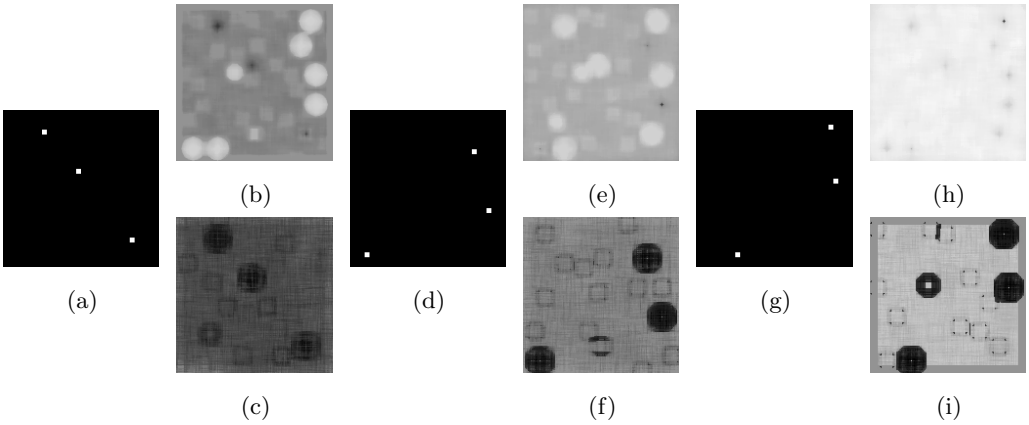


Figure 4.8: Results of the MDPOHMT on a noisy synthetic test image with PO-based relaxation. **a-c)** MDPOHMT, FG dissimilarity, and BG dissimilarity of Figure 4.5d using Figure 4.5m with $p = 82\%$, calculated from Figure 4.5j. **d-f)** MDPOHMT, FG dissimilarity, and BG dissimilarity of Figure 4.5d using Figure 4.5n with $p = 88\%$, calculated from Figure 4.5k. **g-i)** MDPOHMT, FG dissimilarity, and BG dissimilarity of Figure 4.5d using Figure 4.5o with $p = 89\%$, calculated from Figure 4.5l.

Despite each of the objects having varying noise and flipped pixel colours, through analysis of the PO plots and appropriate rank selection they are all successfully detected. It can be noted that the area under the curve of the green (Figure 4.5k) and blue (Figure 4.5l) PO plots is greater than that of the red (Figure 4.5j), this is due to the orange background being more similar to red than either blue or green. As such, the colour of the background SE is an important consideration when designing the composite SE. This is displayed further in Appendix B.1 where similar examples to the one shown in this section are given using multiple background colours (listed in Table B.1) with varying similarity to the red, green, and blue objects of interest.

The rank parameter and selection is intended to be intuitive and represents a percentage fit between an image under test and a probing SE. A percentage is used as there is no need to define which pixels must fit inside the SE. While the PO plot method provides a way of setting the optimal parameter for detecting an object, there may exist some cases where few or no samples are available to train such methods. In these cases, the rank parameter can be estimated easily simply by trialling various percentages, using the understanding that a high percentage fit will restrict the detected objects and a lower percentage will relax the transform [49]. The MDPOHMT is often computationally inexpensive enough to be performed multiple times with different test rank parameters until a positive outcome is reached.

4.5.3 Comparisons with Alternative Morphological Methods

In this section, the MDPOHMT is compared with other HMTs, specifically the UHMT (Section 3.2.5) and SHMT (Section 3.3.3), when detecting objects in various synthetic and natural images. The limitations of these techniques are discussed and the benefits of using the MDPOHMT are highlighted.

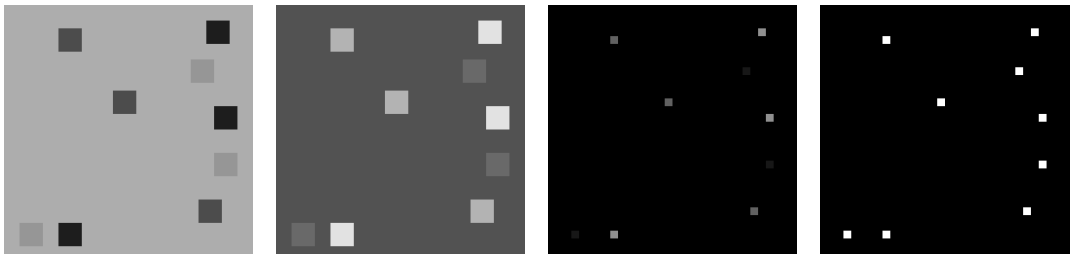
4.5.3.1 Comparison with Soille's Unconstrained HMT

Soille's UHMT, CHMT [27], and other greyscale HMTs have been shown to be useful for various object detection applications [30, 34]. However, there are situations where colour information being taken into account can be advantageous and provide an extra dimension to a decision. To highlight this, the synthetic image from Figure 4.5a was converted to the YIQ colour space using Equation (4.10) and a greyscale image was created from the luminance channel, Y, calculated with Equation (4.11) [272].

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.532 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4.10}$$

$$Y = 0.299R + 0.587G + 0.114B \tag{4.11}$$

This greyscale conversion, and its complement, as well as the result of performing both the UHMT, Equation (3.17), and the general greyscale HMT, Equation (3.19), on the complemented image are show in Figure 4.9.



(a) (b) (c) (d)

Figure 4.9: **a)** Greyscale version of the clean colour test image from Figure 4.4a. **b)** Complement of a). **c)** Result of applying the UHMT to b). **d)** Result of applying the general greyscale HMT to b).

As each of the coloured squares appear darker than the background in Figure 4.9a, none of the coloured foreground objects will be detected by either the UHMT or HMT. Taking the

complement of the image, Figure 4.9b, reverses this, with all nine of the foreground objects becoming brighter than the background. This allows for each square to be detected when using both the UHMT (Figure 4.9c) and HMT (Figure 4.9d) with a greyscale SE. The ability to discriminate between the colour of the desired object is reduced compared to the MDPOHMT result from Figure 4.6. In the UHMT, each colour of foreground object produces a different greyscale value in the output, as a result they can be distinguished as different classes of object but their original colour information is not retrievable. In the HMT, no differentiation of the foreground objects can be achieved. Additional examples of performing the UHMT and greyscale HMT on images with varying background colours compounding these limitations, are shown in Figure B.2.1 (Appendix B.2).

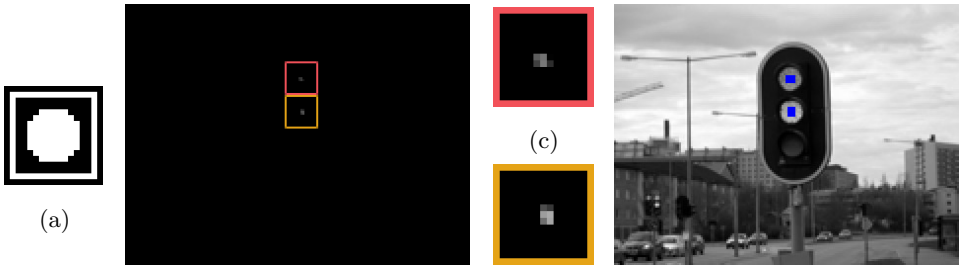
In order to highlight this shortcoming further, an additional experiment was performed comparing the MDPOHMT with the UHMT in object detection performance using an image of a traffic light, where both the red and amber lamps are illuminated. The images used in this comparison, the original RGB image and its greyscale representation obtained using Equation (4.11), are shown in Figures 4.10a and 4.10b respectively.

A composite greyscale SE was created using the size and shape information of both the red and amber illuminated lights in Figure 4.10a and consists of a circular $\boldsymbol{S}_{FG}$ with a radius of 4 pixels and a hollow square $\boldsymbol{S}_{BG}$ with width and height of 15 pixels and thickness of 1 pixel. This SE, and the result of applying the UHMT to the greyscale test image from Figure 4.10b, are shown in Figure 4.11.

Performing the UHMT on the greyscale test image using the greyscale SE from Figure 4.11a yields the results shown in Figure 4.11b. Figures 4.11c and 4.11d provide a detailed view of the locations of the red and amber traffic lights respectively, and show the result of the UHMT in these regions. It can be observed that pixels in the centres of both the red and amber lights

(a) (b)

Figure 4.10: Test image of a traffic light in both **a)** RGB colour and **b)** YIQ luminance



(a) (b) (c) (d) (e)

Figure 4.11: **a)** Circular SE used to detect lights. **b)** Result of performing the greyscale UHMT on Figure 4.10b using a). **c)** UHMT values in the region of the red light from b). **d)** UHMT values in the region of the amber light from b). **e)** Result of taking the UHMT overlaid onto the test image (dilated by a 3×3 SE).

are identified. This is due to the SE fitting both objects and, similarly to the example from Figure 4.9, there is limited ability to distinguish between the original colours. The range of UHMT output values are 0.17 – 0.47 and 0.22 – 0.66 for identified pixels in the red, Figure 4.11c, and amber, Figure 4.11d, lights respectively, meaning that a simple threshold is insufficient to delineate them. Context, such as spatial location or contiguity may be used to separate objects, but they cannot then be classified based on their original colour. This highlights the benefits of incorporating colour information into the HMT in order to better discriminate between objects whose colours have similar greyscale intensities where size and shape information is not a distinguishing feature.

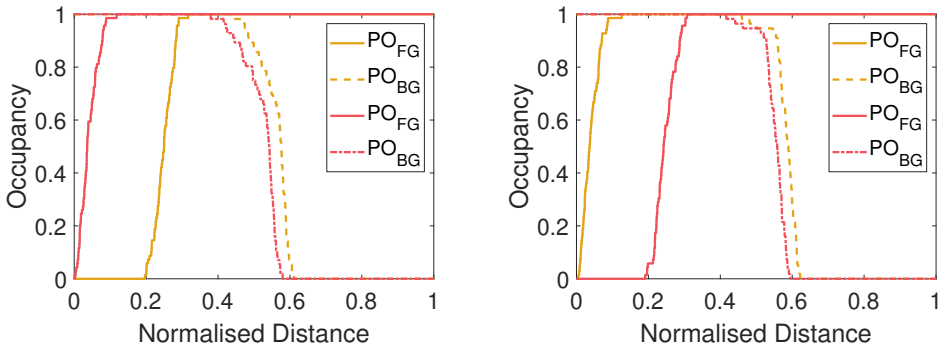
The average colour within each of the red and amber traffic lights was used as the colour of both $\boldsymbol{S}_{FG}$ and $\boldsymbol{S}_{BG}$ for use with the MDPOHMT. Combining this with the size and shape information captured in the greyscale SE from Figure 4.11a produces the two RGB colour SEs shown in Figures 4.12a and 4.12b respectively. The results of performing the MDPOHMT on the image from Figure 4.10a using these red and amber composite SEs are shown in Figures 4.12c and 4.12d respectively.

Initially, with a PO value of $p = 100\%$, where $p_{FG} = p_{BG} = p$, and a threshold of $T = 1$, as with the UHMT, the MDPOHMT is unable to distinguish between the red and amber lights when using either the red or amber SEs. By analysing the occupancy graphs and PO plots for the centre pixel of each light, shown in Figures 4.13 and 4.14 respectively, this behaviour can be explained.

Figures 4.13a and 4.13b show the occupancy of both the red and amber SEs when translated to the centre of the red and amber lights respectively. In both cases, there are a range of distances where 100% of the pixels within the support of $\boldsymbol{S}_{FG}$ are less than the distance to the image while simultaneously 100% of the pixels within the support of $\boldsymbol{S}_{BG}$ are further than the

(a) (b) (c) (d)

Figure 4.12: **a)** Composite SE used to find the red light. **b)** Composite SE used to find the amber light. **c)** Result of performing the MDPOHMT on Figure 4.10a using a) with $p = 100\%$. **d)** Result of performing the MDPOHMT on Figure 4.10a using b) with $p = 100\%$.



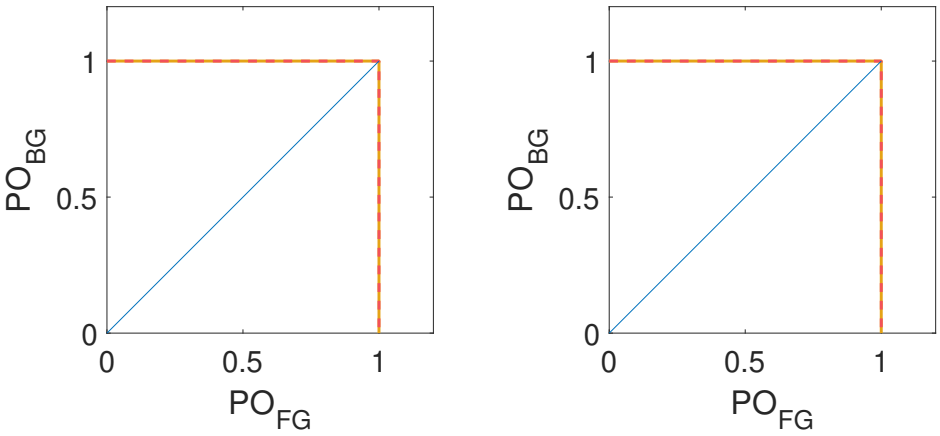
(a) (b)

Figure 4.13: Occupancy of the red and amber SEs when translated to the centre of the **a)** red and **b)** amber lights

same distance to the image. As a result, both lights are detected regardless of the SE used. The corresponding PO plots, Figures 4.14a and 4.14b for the red and amber lights respectively, confirm both lights can be detected without any relaxation.

There are two approaches to improve the ability of the MDPOHMT to discriminate between objects of similar colours; change the design of the SE, or impose more stringent requirements for a match. Figure 4.15 shows the results from the first of these two options, where $\boldsymbol{S}_{BG}$ in each composite colour SE, $\boldsymbol{S}$, has been changed from red and amber respectively to a grey colour, as shown in Figures 4.15a and 4.15b. This has the effect of decreasing the distance of $\boldsymbol{S}_{BG}$ to the image background in areas where a match is expected (*i.e.,* the black traffic light housing). The results of applying the MDPOHMT with this updated SE design are shown in Figures 4.15c and 4.15d using Figures 4.15a and 4.15b respectively.

As shown in Figures 4.15c and 4.15d, these updated SEs have the desired effect and are able to detect the corresponding object in the image. Investigating the updated PO plots produced



(a) (b)

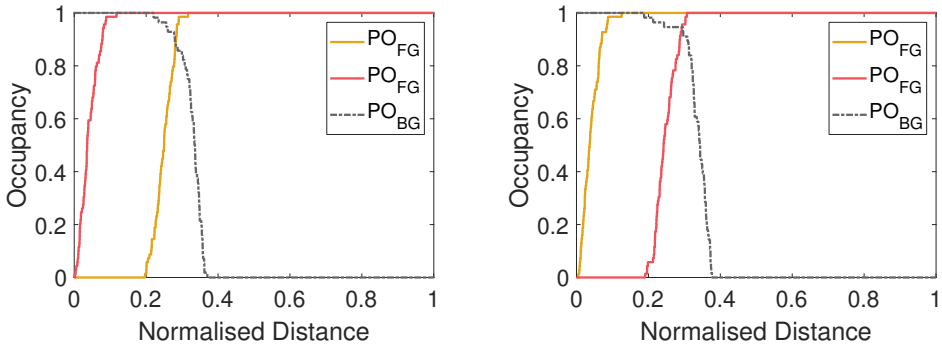
Figure 4.14: PO plots of the red and amber SEs when translated to the centre of the **a)** red and **b)** amber lights

(a) (b) (c) (d)

Figure 4.15: **a)** Updated circular SE used to find the red light. **b)** Updated circular SE used to find the amber light. **c)** Result of performing the MDPOHMT on Figure 4.10a using a) with $p = 100\%$. **d)** Result of performing the MDPOHMT on Figure 4.10a using b) with $p = 100\%$.

when each of these new SEs are translated to the centres of each traffic light, as shown in Figures 4.16 and 4.17, provides some insight into this new result.

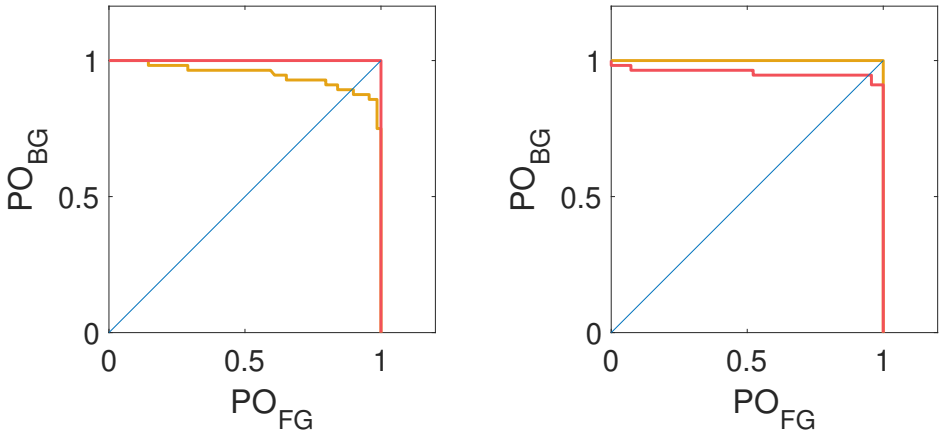
Comparing the occupancy plots from Figure 4.16, produced using the updated $\boldsymbol{S}_{BG}$, with those from Figure 4.13, using the original, the background occupancy, or the proportion of pixels in the support of $\boldsymbol{S}_{BG}$ greater than some dissimilarity, reduces to zero at lower dissimilarities as it has been made more similar to the objects' background. As a result, when centred at the light it has not been designed for, the PO value required for a match is decreased, *e.g.,* in Figure 4.16a where the background occupancy vector has shifted leftwards, bringing the point of intersection between PO_{FG} for the amber light and PO_{BG} downwards. An identical effect can be seen in Figure 4.16b for the red light. This results in each SE detecting the intended light, and no longer producing a match at 100% occupancy when the amber SE is centred on



(a) (b)

Figure 4.16: Occupancy of the updated red and amber SEs when translated to the centre of the **a)** red and **b)** amber lights

the red light and vice versa. This lack of a perfect match, *i.e.,* a right angle for both SEs, is shown in Figures 4.17a and 4.17b where the only the appropriate SE for a given object matches.

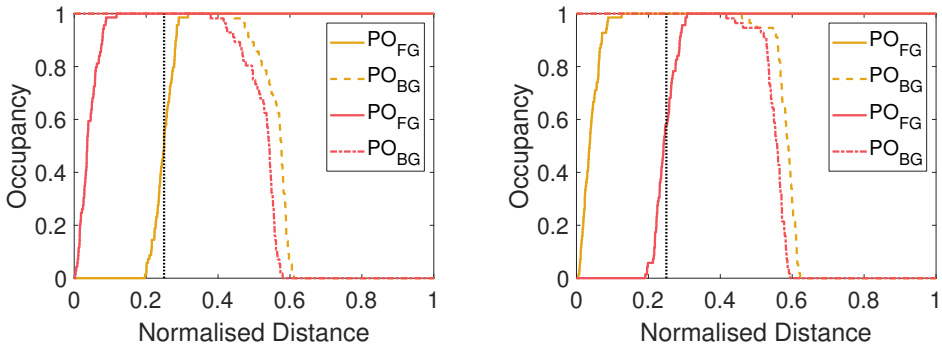


(a) (b)

Figure 4.17: PO plots of the updated red and amber SEs when translated to the centre of the **a)** red and **b)** amber lights

The second option, instead of altering the SE design, is to impose stricter limits for what is deemed to be a match. As both objects are detected with a PO of 100%, this cannot be limited further. However, the foreground dissimilarity threshold, T, can be used in such situations to isolate objects of interest. Investigating the occupancy graph for the pixel at the centre of the red light, there is a range of distances where, at $p = 100\%$, the red foreground SE is at 100% occupancy and the amber foreground SE, centred on the same point, is not. By specifying that the foreground dissimilarity must be less than or equal to some value in this range, the amber foreground SE is no longer a close enough match to detect the red light. An identical process can be followed to exclude the red SE when centred on the amber light. This

is shown in Figure 4.18, where a black dotted vertical line represents a threshold at $T = 0.25$ in each of the occupancy graphs for the red and amber lights. The results of applying the MDPOHMT with the original colour composite SEs with a reduced foreground threshold are shown in Figure 4.19. Reducing the foreground dissimilarity threshold has the desired effect of preventing the non-matching SE from detecting the other light, with the red SE detecting the red light and the amber SE detecting solely the amber light. This highlights the ability of the MDPOHMT to distinguish between objects of interest based on their colour information as well as size and shape in natural colour images. It should be noted, however, that decreasing the value of p may require also decreasing the value of T. An example of this is shown in Figure 4.18b, where the foreground occupancy of the red SE intersects the vertical threshold drawn at $T = 0.25$ at an occupancy of $p \approx 55\%$ and would be detected using these parameters.



(a) (b)

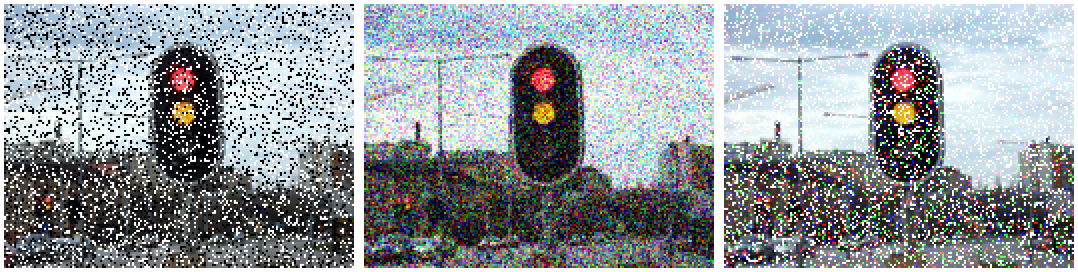
Figure 4.18: Occupancy of the red and amber SEs when translated to the centre of the **a)** red and **b)** amber lights with the addition of a colour threshold, T.

(a) (b) (c) (d)

Figure 4.19: **a)** Circular SE used to find the red light. **b)** Circular SE used to find the amber light. **c)** Result of performing the MDPOHMT on Figure 4.10a using a) with $p = 100\%$ and $T = 0.25$. **d)** Result of performing the MDPOHMT on Figure 4.10a using b) with $p = 100\%$ and $T = 0.25$.

4.5.3.2 Effect of Noise on the MDPOHMT

The use of PO as a form of noise robustness in the MDPOHMT is necessary to detect objects in non-ideal imagery, such as in the presence of noise or where the objects of interest have been occluded. In order to test its robustness in various noise conditions, three noise models were used; bipolar impulsive (salt-and-pepper) noise, Gaussian noise, and impulsive (salt) noise. Each noise model is defined based on a strength. In the case of the impulsive noise models, this is defined as the probability of noise, $\mathrm{P}(n)$, whereas for Gaussian noise, it is the noise variance, σ^2. In addition, a channel correlation parameter $0 \leq \rho \leq 1$ is used such that if noise occurs in a certain pixel, there is a probability of ρ that this noise is applied to all of the channels, and a probability of $(1-\rho)/N$ that it is applied to any of the N channels in the image individually. An example of each of these noise models being applied to the test image of a traffic light from Figure 4.10a is shown in Figure 4.20.



(a) (b) (c)

Figure 4.20: Example noise models used to test the MDPOHMT. **a)** Bipolar Impulsive noise $\mathrm{P}(n) = 0.3$, $\rho = 1$. **b)** Gaussian noise $\sigma^2 = 0.05$, $\rho = 0$. **c)** Impulsive noise $\mathrm{P}(n) = 0.3$, $\rho = 0.5$.

The bipolar impulsive noise applied to the image in Figure 4.20a affects approximately 30% of the image pixels, where the affected pixels are set to black or white due to the channel correlation value $\rho = 1$, *i.e.,* where noise exists, each of the channels' values are set to either 0 or 1. The other impulsive noise model, shown in Figure 4.20c, also affects approximately 30% of pixels, however, half of these will be set to white, and the others, affecting only a single channel, result in pure red, green, or blue pixels, accounting for roughly 5% of the total pixels each. The Gaussian noise in Figure 4.20b affects the entire image, with different layers of noise added to each channel individually as $\rho = 0$.

In order to test the MDPOHMT in varying scenarios, multiple combinations of noise models and noise parameters were applied to the image from Figure 4.10a. The set of noise strength parameters used were $\mathrm{P}(n) = \{0.1,\ 0.3\}$ and $\sigma^2 = \{0.01,\ 0.05\}$ for the impulsive and Gaussian noise models respectively, and the set of channel correlations were $\rho = \{0,\ 0.5,\ 1\}$. PO plots

for each type of noise model and combination of parameters are shown in Figure 4.21 for both the red and amber lights.

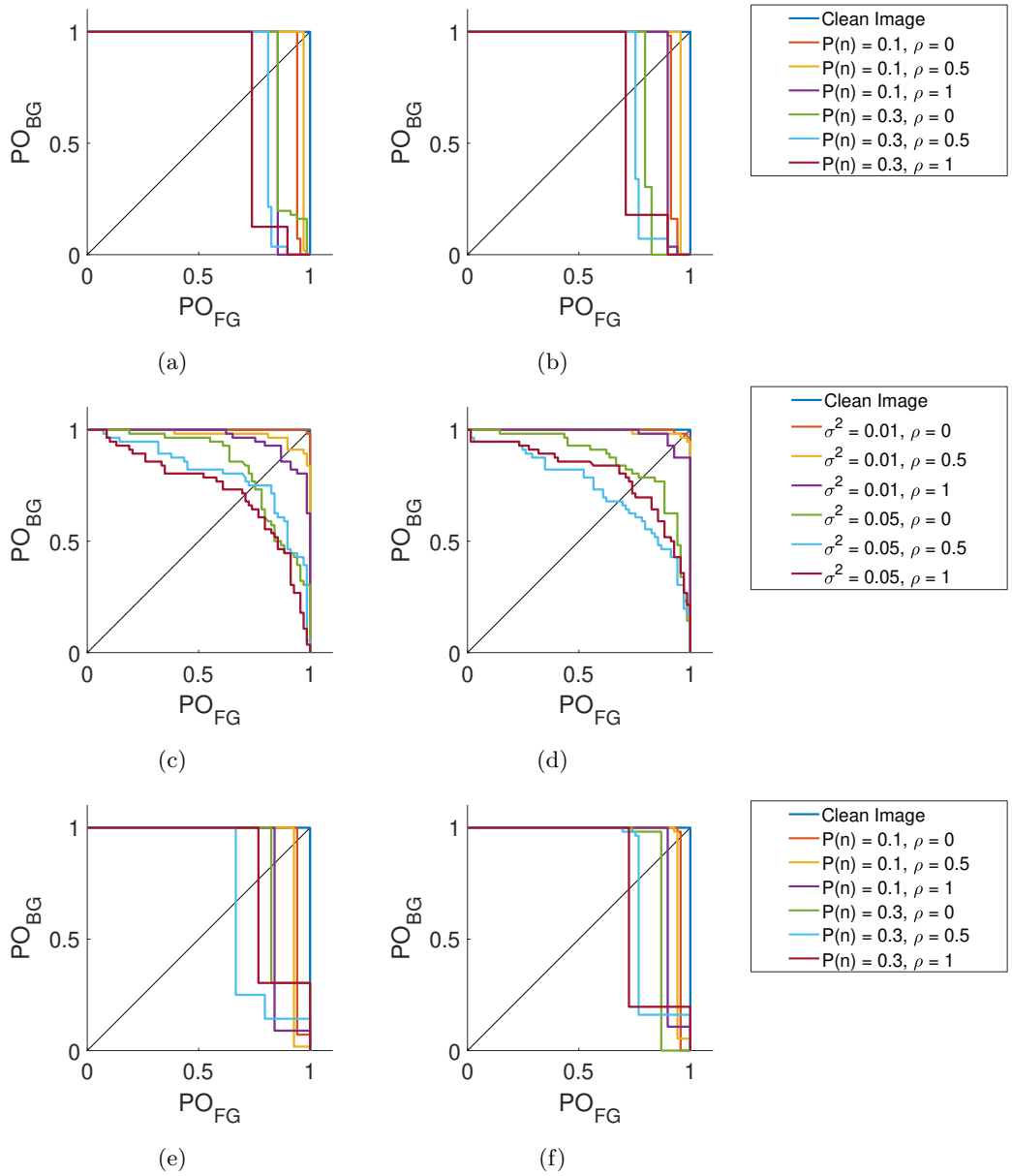
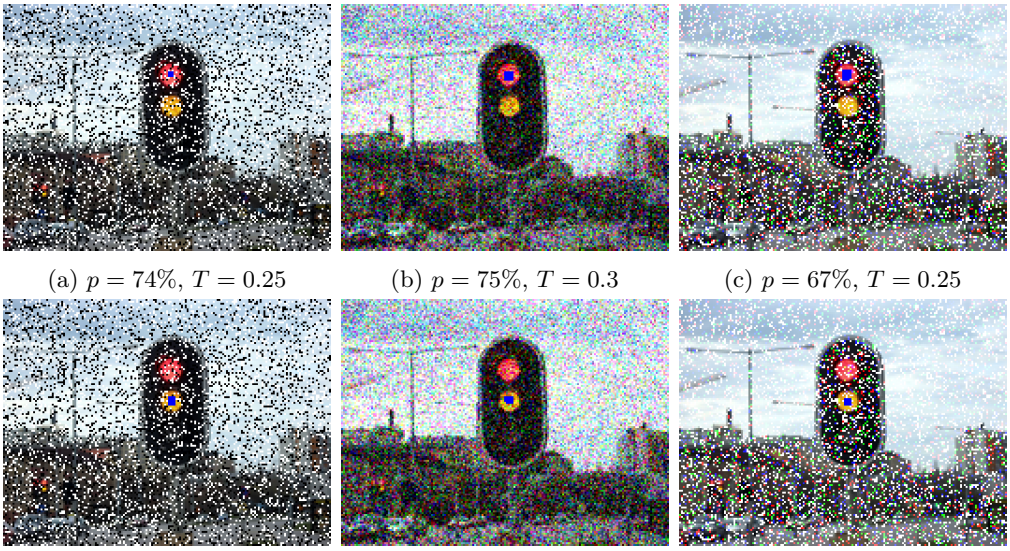


Figure 4.21: PO plots at the centre pixels of both traffic lights in the noise corrupted for each combination of noise parameters. PO plot of the **a)** red and **b)** amber lights under each combination of Bipolar Impulsive noise. PO plot of the **c)** red and **d)** amber lights under each combination of Gaussian noise. PO plot of the **e)** red and **f)** amber lights under each combination of Impulsive noise.

The results of detecting the red and amber lights in the subset of noise corrupted images from Figure 4.20 using the MDPOHMT are shown in Figure 4.22. The MDPOHMT parameters

used are given below each of the detections presented in Figure 4.22. The PO value, p, used was calculated using the PO plots in Figure 4.21 and the foreground threshold $T = 0.25$ was identical to that calculated in Figure 4.18, with the exception of the images affected by Gaussian noise where it was increased to $T = 0.3$. The full set of results and the parameters used on each of the noisy images are shown in Appendix B.3. In each case for all of the combinations of noise models and parameters tested, both the red and amber lights were detected without false positive detections when using the corresponding SE and calculating the required rank from the PO plots.



(a) $p = 74\%$, $T = 0.25$ (b) $p = 75\%$, $T = 0.3$ (c) $p = 67\%$, $T = 0.25$

(d) $p = 71\%$, $T = 0.25$ (e) $p = 79\%$, $T = 0.3$ (f) $p = 77\%$, $T = 0.25$

Figure 4.22: Results of applying the MDPOHMT on noise corrupted images. Detecting the red and amber light in the presence of; **a,d)** Bipolar impulsive noise, **b,e)** Gaussian noise, and **c,f)** Impulsive noise.

4.5.3.3 Comparison with the Similarity HMT

In this section, the MDPOHMT is compared with another similar colour HMT approach, the SHMT [40]. The MDPOHMT is based on a reduced ordering between a multivariate image and a multivariate SE, transforming this multivariate data into a scalar for ordering. The SHMT calculates the colour similarity between a colour image and SE, as described in Section 3.3.3 and defined in Equation (3.33). These similarity values along with the image's individual colour channels are ordered lexicographically in similarity, then green, then red, then blue order as shown in Equation (3.32). The infimum, or erosion, of the ordered similarity within the support of a probing SE is taken and the SHMT is defined as all points in the image where the resultant value of similarity is equal to one. The SHMT$_\eta$, Equation (3.34), relaxes this

and is defined as all points in an image where the minimum similarity in the SE window is greater than some threshold η. This differs from the PO-based noise robustness offered by the MDPOHMT developed in this thesis. A comparison between the MDPOHMT and the SHMT on the synthetic test images presented in Figure 4.4 is shown in Figure 4.23.

Figures 4.23a to 4.23c show the clean, flipped, and noisy synthetic colour test images from Figure 4.5. These images are used compare the behaviours of the SHMT and the MDPOHMT. As the SHMT only requires a single foreground SE, the red $\boldsymbol{S}_{FG}$ from Figure 4.5m is used. Figures 4.23d to 4.23f show the pixel-wise similarity as a result of translating $\boldsymbol{S}_{FG}$ to each point in the image and calculating the resultant erosion of the set of similarities after they have been ordered using Equation (3.32). This similarity image is essentially a visualisation of the least similar pixels in the image within the support of a probing SE and is similar to the foreground distance images produced by the MDPOHMT, *e.g.,* those shown in Figure 4.6. It can be observed that when $\boldsymbol{S}_{FG}$ is translated to each of the red objects in the clean image there is a single pixel of similarity equal to 1 in the centre of each object, as shown in Figure 4.23d, whereas in Figures 4.23e and 4.23f there are large regions which exhibit low, or zero, similarity between the image and $\boldsymbol{S}_{FG}$.

The result of performing the SHMT on each of the images in Figures 4.23a to 4.23c, essentially the set of pixels equal to 1 in the similarity images from Figures 4.23d to 4.23f, are shown in Figures 4.23g to 4.23i where they have been dilated for ease of viewing. As expected, all three of the red objects are detected in the clean image using the SHMT, with only a single object, corresponding to the one with no flipped pixels, detected in the flipped image, and zero detected in the noisy image. This is attributable to the fact that in the similarity images for both of these images (Figures 4.23e and 4.23f) there are fewer pixels with a prefect similarity.

In addition to the SHMT, the noise robust variant, the SHMT_η, was tested on both the flipped and noisy images, to investigate the performance of the SHMT in the presence of noise and non-ideal objects. The SHMT_η applies a single threshold, η, across the similarity image in order to relax the similarity required for the detection of an object given some degradation. The similarity of each pixel in Figures 4.23d to 4.23f is in the range $0 \leq h \leq 1$. By analysing the centre pixels of each red object of interest in the similarity images, the observed similarities in the clean, flipped, and noisy images, respectively, are $\{1,\ 1,\ 1\}$, $\{1,\ 0.63,\ 0\}$, and $\{0.61,\ 0.46,\ 0\}$. With these similarities in mind, performing the SHMT_η, where $\eta = 0.6$ *i.e.,* $\text{SHMT}_{0.6}$, should be sufficient to detect two of the objects in the flipped image (Figure 4.23b) and one in the noisy image (Figure 4.23c). The results of performing the $\text{SHMT}_{0.6}$, on the flipped and noisy images are shown in Figures 4.23j and 4.23k respectively. Additionally, a similarity of $\eta = 0.4$

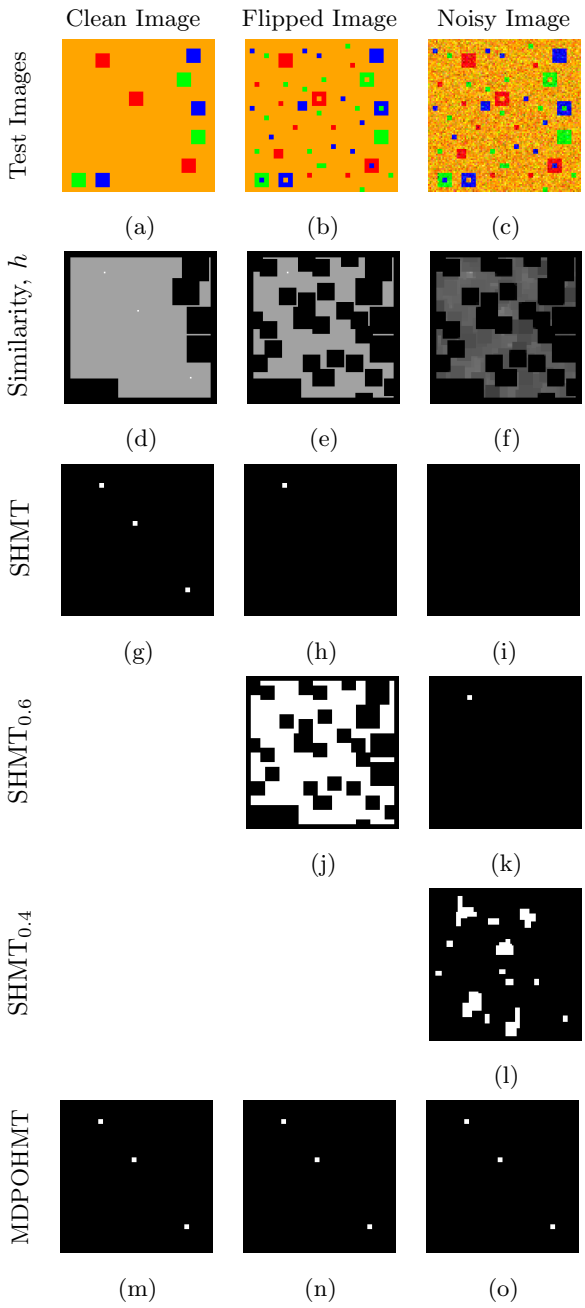


Figure 4.23: Comparison between the SHMT and MDPOHMT on synthetic data. **a-c)** Test images used. **d-f)** Similarity between the red $\boldsymbol{S}_{FG}$ in Figure 4.5m and the images in a-c). **g-i)** SHMT result. **j,k)** $\text{SHMT}_{0.6}$ result on the flipped and noisy images. **l)** $\text{SHMT}_{0.4}$ result on the noisy image. **m)** MDPOHMT with $p = 1$ and $T = 0.2$. **n)** MDPOHMT with $p = 0.8$ and $T = 0.2$. **o)** MDPOHMT with $p = 0.7$ and $T = 0.2$.

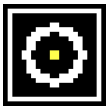
is sufficient to detect the middle red object in the noisy image, so η is reduced further and the result of the $\text{SHMT}_{0.4}$ can be seen in Figure 4.23l.

As shown in Figure 4.23j, when reducing the similarity threshold to $\eta = 0.6$ to detect the middle red object containing the orange flipped pixels, the majority of the orange background is also detected. This is due to the minimum similarity within the support of the red SE being identical when centred over the object or the background. The same value of η is sufficient to detect the red object in the top left of the noisy image in Figure 4.23c as it has a similarity of 0.61. The result of reducing η further to detect additional objects in the noisy image is shown in Figure 4.23l. While this reduction in η has the desired effect of adding the centre pixel of the middle red object to the set of pixels detected by the $\text{SHMT}_{0.4}$, there are a number of FP detections introduced as a result. It should also be noted that the object in the lower right corner cannot be detected using the SHMT_{η} in either the flipped or noisy image due to there being zero chromatic similarity between blue and red and, as the SHMT is based on the erosion, or minimum, similarity within the support of a SE, this will always produce approximately zero similarity, dependent on noise, when intersecting with a probing red SE.

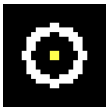
The SHMT, and by extension the SHMT_{η}, is essentially operating as an erosion and as such only considers the minimal value of similarity between the image pixels and corresponding reference values within the support of an SE. This could be improved by using a rank order filter to relax this erosion in a similar manner to the proposed MDPOHMT. If the hard minimum similarity is relaxed, situations like those presented in Figures 4.23b and 4.23c, where there are low similarity pixels present in the support of an SE causing the SHMT to miss an object can be avoided. By incorporating such rank order filters the MDPOHMT can be made more robust to these effects of noise. The results of the MDPOHMT with appropriate PO values set from Figures 4.5b, 4.5f and 4.5j respectively, where each of the objects are identified perfectly, with no false positives are shown in Figures 4.23m to 4.23o for reference. This highlights the advantages of using a PO-based noise relaxation, as the value of (dis)similarity returned is reduced, rather than just a threshold.

A further comparison between the performance of the MDPOHMT and the SHMT when applied to a natural image was then carried out. Figure 4.24c shows an image of white and yellow flowers in a grassy background. This image was used in [217] to validate the performance of the SHMT and SHMT_{η}. Figures 4.24a and 4.24b show the SEs used in the both the MDPOHMT and SHMT respectively, where Figure 4.24b is the foreground SE of the composite SE in Figure 4.24a. $\boldsymbol{S}_{FG}$ consists of a 2×2 yellow square in the centre surrounded by a white annulus of radius 6 and thickness 2. This is designed to act as the skeleton of the objects of

interest. $\boldsymbol{S}_{BG}$ is a hollow 20×20 white square with a thickness of 1 pixel, designed to fit around each of the individual flowers.



(a)



(b)

(c)

Figure 4.24: **a)** Composite SE used in the MDPOHMT. **b)** SE used in the SHMT. **c)** "*Flowers*" test image [217].

The similarity value obtained by translating the SE from Figure 4.24b to each pixel of Figure 4.24c is shown in Figure 4.25a. The majority of the image is black or dark grey, meaning there is a low minimum similarity between the image and SE at those pixels. The SHMT_η is essentially a thresholding operation on this similarity image where η represents the value of the threshold. Reducing η allows for the detection of decreasingly similar objects. The results of carrying out the SHMT_η, where η is reduced in each subsequent test: $\eta = \{1,\ 0.9,\ 0.8,\ 0.7,\ 0.6,\ 0.5,\ 0.4,\ 0.3\}$, are shown in Figures 4.25b to 4.25i. The SHMT_η operator does not allow for any relaxation of the shape of an SE and therefore requires a much more specific SE design and is sensitive to variation in the shape of objects of interest as well as noise or occlusion.

The maximum value for similarity observed in Figure 4.25a was 0.93, meaning that performing the SHMT without any relaxation did not result in any pixels being identified as belonging to a flower, as shown in Figure 4.25b. Reducing the threshold to $\eta = 0.9$ however, results in the centres of two flowers, denoted by red pixels overlaid onto the image, being detected in Figure 4.25c. These two detected objects are the most similar of the objects in the image to the SE and are free from any damage or occlusion that would otherwise lower their minimum similarity and prevent their detection. Lowering the similarity threshold to $\eta = \{0.8,\ 0.7,\ 0.6\}$ results in the images shown in Figures 4.25d to 4.25f, where additional flowers are detected and the regions identified as matching grow.

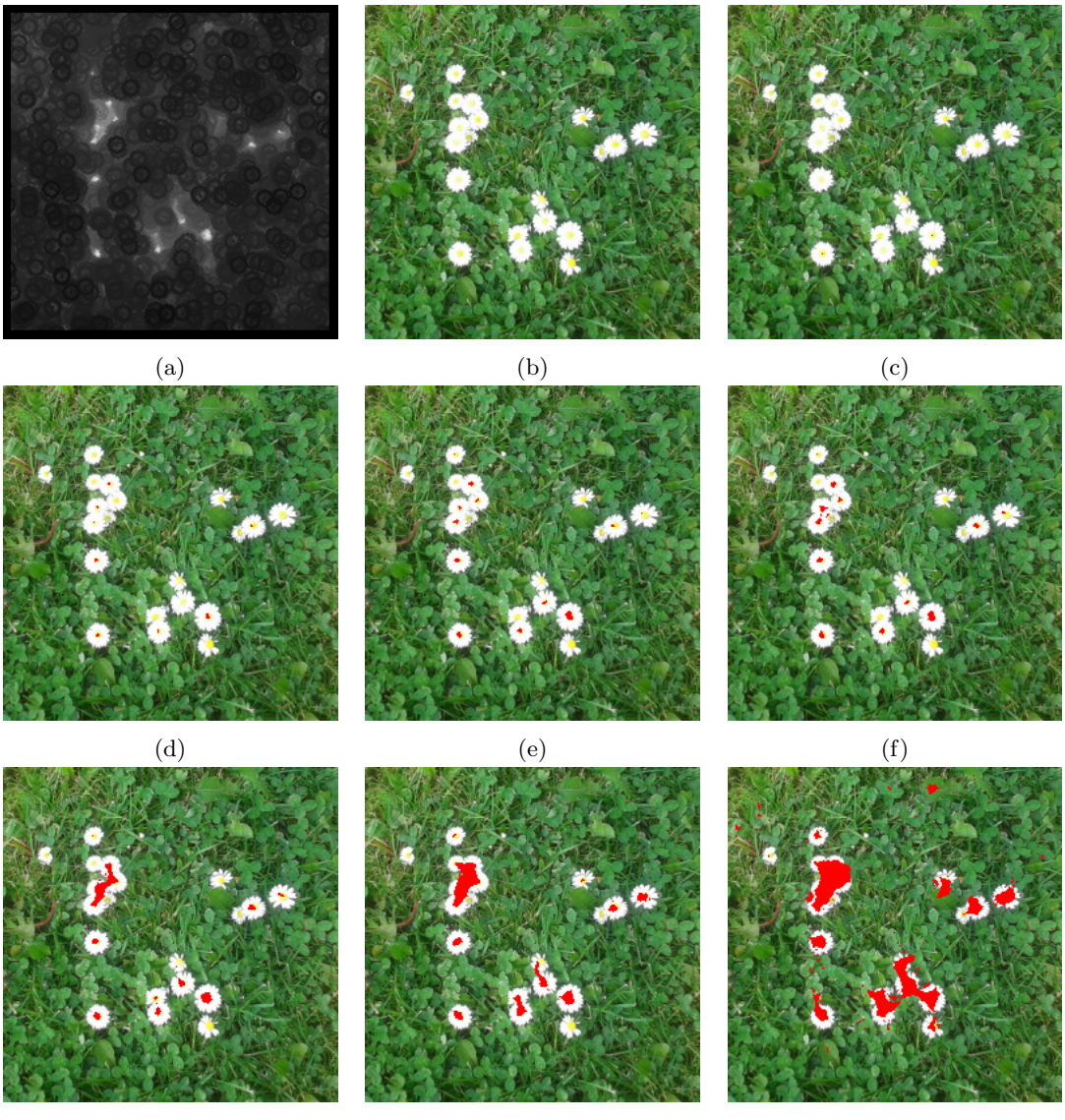


Figure 4.25: SHMT results on the "*Flowers*" test image. **a)** Similarity image between Figure 4.24b and Figure 4.24c. **b)** SHMT. **c)** $\text{SHMT}_{0.9}$. **d)** $\text{SHMT}_{0.8}$. **e)** $\text{SHMT}_{0.7}$. **f)** $\text{SHMT}_{0.6}$. **g)** $\text{SHMT}_{0.5}$. **h)** $\text{SHMT}_{0.4}$. **i)** $\text{SHMT}_{0.3}$.

Reducing the threshold further to $\eta = \{0.5,\ 0.4,\ 0.3\}$, produces the results shown in Figures 4.25g to 4.25i. In Figure 4.25g, over half of the flowers present in the scene have at least one positively detected pixel. However, a patch of positively identified pixels, spread across a number of closely grouped flowers, is present. This patch is not representative of the object that the SE, Figure 4.24b, has been designed to detect, *i.e.,* the yellow centre of these flowers. This occurs because η is set to a level which allows pixels with relatively low minimum similarity to be included in the set of detections. Referring to Figure 4.25a, which shows the similarity of each pixel to the probing SE, it is possible to see that there are some flowers which exhibit low similarity. As a result, in order to detect them, η must be set at such a low level that a number of false positive are introduced, as is the case in Figure 4.25i which is similar to the synthetic result in Figures 4.23j and 4.23l. These similarities for a centre pixel in each of the flowers are given in Table 4.1. In [217], the author explains that some flowers cannot be detected as they are smaller than the generalised SE or that they are damaged and, as a result, exhibit low minimum similarity to the SE. This limitation is countered by the notion of PO, and by extension the MDPOHMT.

In order to estimate the PO values required to detect each of the flowers, the occupancy and PO plots for a subset of centre pixels were generated, as shown in Figures 4.26b and 4.26c, corresponding to the locations marked in Figure 4.26a. The estimated PO values for the centre pixel for each of the flowers, as well as their similarity value from Figure 4.25a, are listed in Table 4.1.

Table 4.1: Location, PO, and Similarity of the centre pixel of each flower in the test image (Figure 4.24c).

ID	X	Y	Estimated PO	Similarity	ID	X	Y	Estimated PO	Similarity
1	**70**	**54**	**0.87**	**0.34**	⋮	⋮	⋮	⋮	⋮
2	**34**	**68**	**0.76**	**0.29**	**11**	**191**	**108**	**0.96**	**0.82**
3	70	74	0.78	0.46	12	181	113	0.74	0.31
4	82	76	0.78	0.82	13	72	134	0.97	0.85
5	**168**	**87**	**0.78**	**0.40**	14	135	149	0.72	0.32
6	87	89	0.71	0.40	**15**	**138**	**165**	**0.84**	**0.63**
7	73	95	0.75	0.77	16	116	177	0.83	0.51
8	213	99	0.92	0.56	17	158	177	0.82	0.90
9	**80**	**102**	**0.53**	**0.27**	18	120	189	0.88	0.80
10	71	106	0.84	0.74	19	73	191	1.00	0.93
⋮	⋮	⋮	⋮	⋮	**20**	**158**	**198**	**0.86**	**0.33**

The lines in bold correspond to those used in the PO plots shown in Figure 4.26.

From Figure 4.26, it is possible to determine that each of the flower centre pixels investigated, with the exception of one corresponding to ID 9 in Table 4.1, can be found with $p \approx 70\%$ occupancy. The results of performing the MDPOHMT with $p_{FG} = \{0.9,\ 0.8,\ 0.7,\ 0.6\}$,

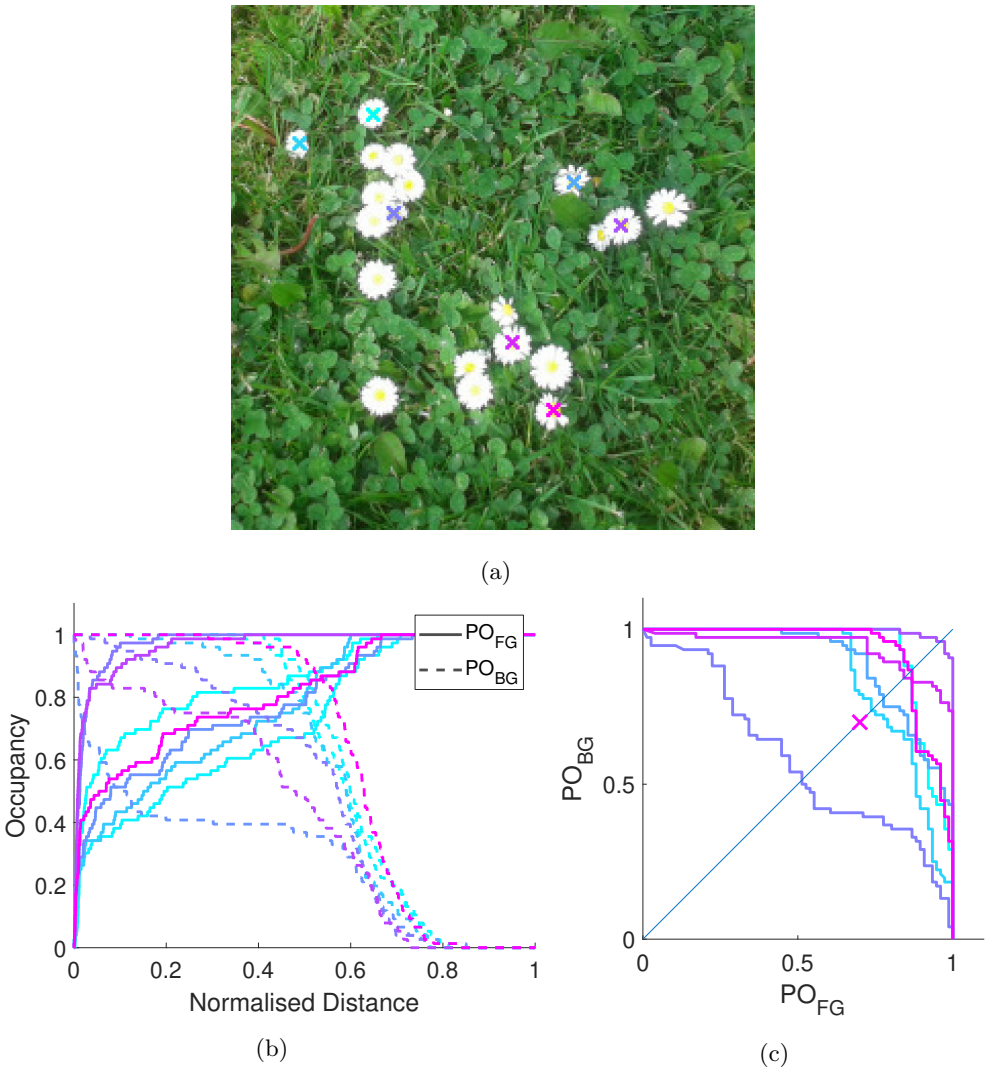


Figure 4.26: **a)** Locations of the subset of objects for which PO plots were generated. **b)** Occupancy plot for the locations in a). **c)** PO plot at the locations in a).

$p_{BG} = 0.75$, and $T = 0.4$ are shown in Figures 4.27b, 4.27d, 4.27f and 4.27h, along with the corresponding foreground distance images, which serve as a comparison to the similarity image from Figure 4.25a, shown in Figures 4.27a, 4.27c, 4.27e and 4.27g respectively.

The MDPOHMT outperforms the SHMT even with high values of PO, *e.g.,* at $p = 90\%$, Figure 4.27b, it detects the majority of the flowers detected using the $\text{SHMT}_{0.5}$. Additionally, it is much more precise with no false positive detections, as seen when comparing the MDPOHMT result with that of the SHMT, *i.e.,* Figures 4.25g and 4.27b.

In order to detect the other damaged or smaller flowers, the foreground PO value was reduced in steps to $p = 60\%$ as shown in Figures 4.27d, 4.27f and 4.27h. This improves the number of flower centres detected but also introduces some patches of white petals into the set of detected pixels. However, unlike the SHMT, the inclusion of $\boldsymbol{S}_{BG}$ regulates this by maintaining that a set percentage, in this case 75%, had to be non-intersecting with the white petals. As the value of p_{FG} decreases, particularly where $p_{FG} = 0.6$ as shown in Figure 4.27h, areas where white petals are prevalent are also detected as they are not dissimilar to the SE. However, by lowering PO the similarity required for a fit is not decreased, as this is regulated by the threshold, T, and instead it is merely an allowance for dissimilar pixels to exist within the support of $\boldsymbol{S}$ and be ignored.

As discussed in Section 2.2.2 and shown in Figures 4.23 and 4.25, the SHMT is susceptible to a single noisy pixel with low chromatic similarity to the target requiring such a low threshold to be set that large numbers of false positives are introduced. The h-similarity measure may be incorporated as a distance measure within the MDPOHMT and by relaxing the percentage fit using rank order filters, as well as varying the thresholded value, η, the detection results can be improved when compared to using the SHMT_{η}. Replacing the Euclidean distance with the h-similarity measure used in the SHMT, and taking advantage of the PO-based relaxation provided by the MDPOHMT, yields the results shown in Figure 4.28. To treat the similarity, h, as a normalised distance/dissimilarity measure, as required by the MDPOHMT, the complement is taken, *i.e.,* $1 - h$. As a result, using similarity with the MDPOHMT, T becomes equivalent to η. In both Figure 4.27 using the Euclidean distance, and Figure 4.28 for the h-similarity, the foreground distance images get slightly darker as the PO value decreases compared to the static similarity image in Figure 4.25a. This relaxation of the foreground dilation and background erosion is one of the main advantages of the MDPOHMT over similar methods such as the SHMT. Since dissimilar pixels within the support of the SE can be ignored, the MDPOHMT can achieve greater noise robustness and improved detection of occluded, damaged, or variable objects in an image.

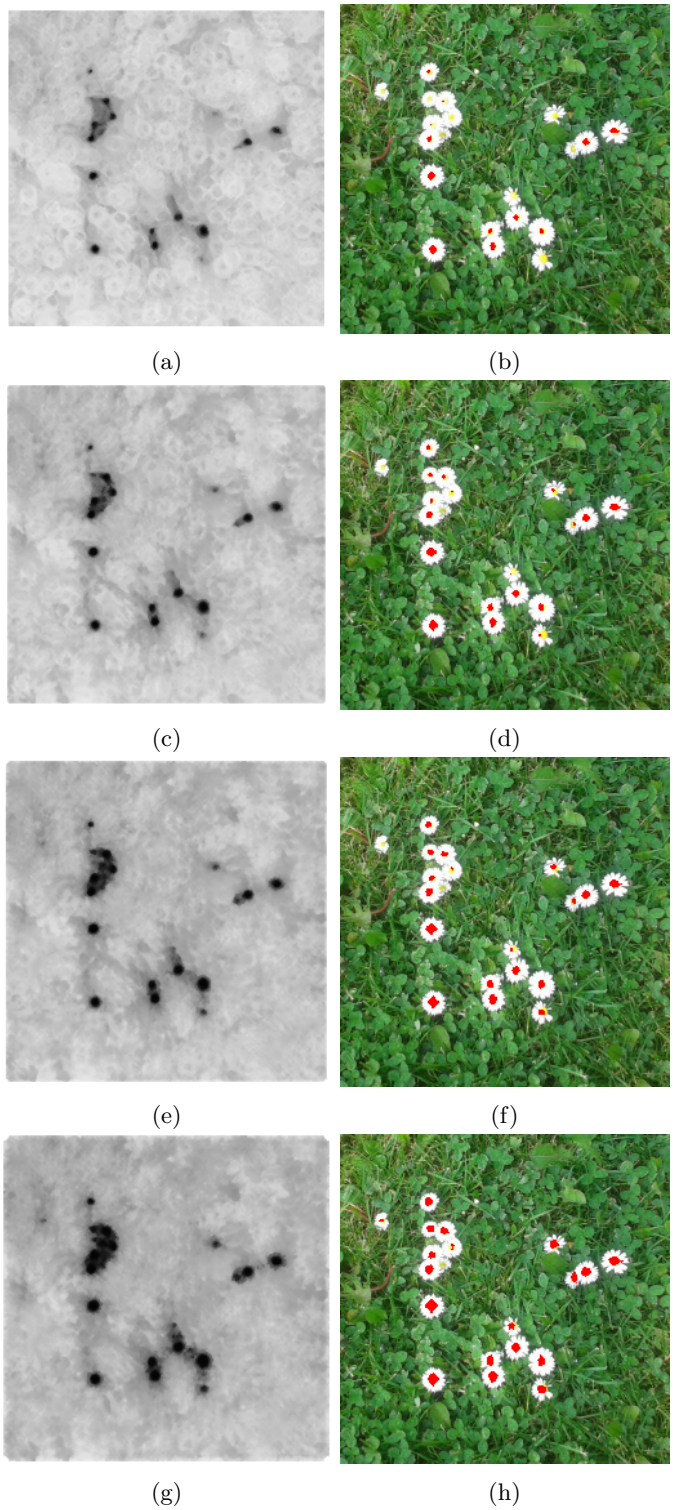


Figure 4.27: Foreground distance image and MDPOHMT result using the Euclidean distance with **a,b)** $p = 90\%$, **c,d)** $p = 80\%$, **e,f)** $p = 70\%$, **g,h)** $p = 60\%$.

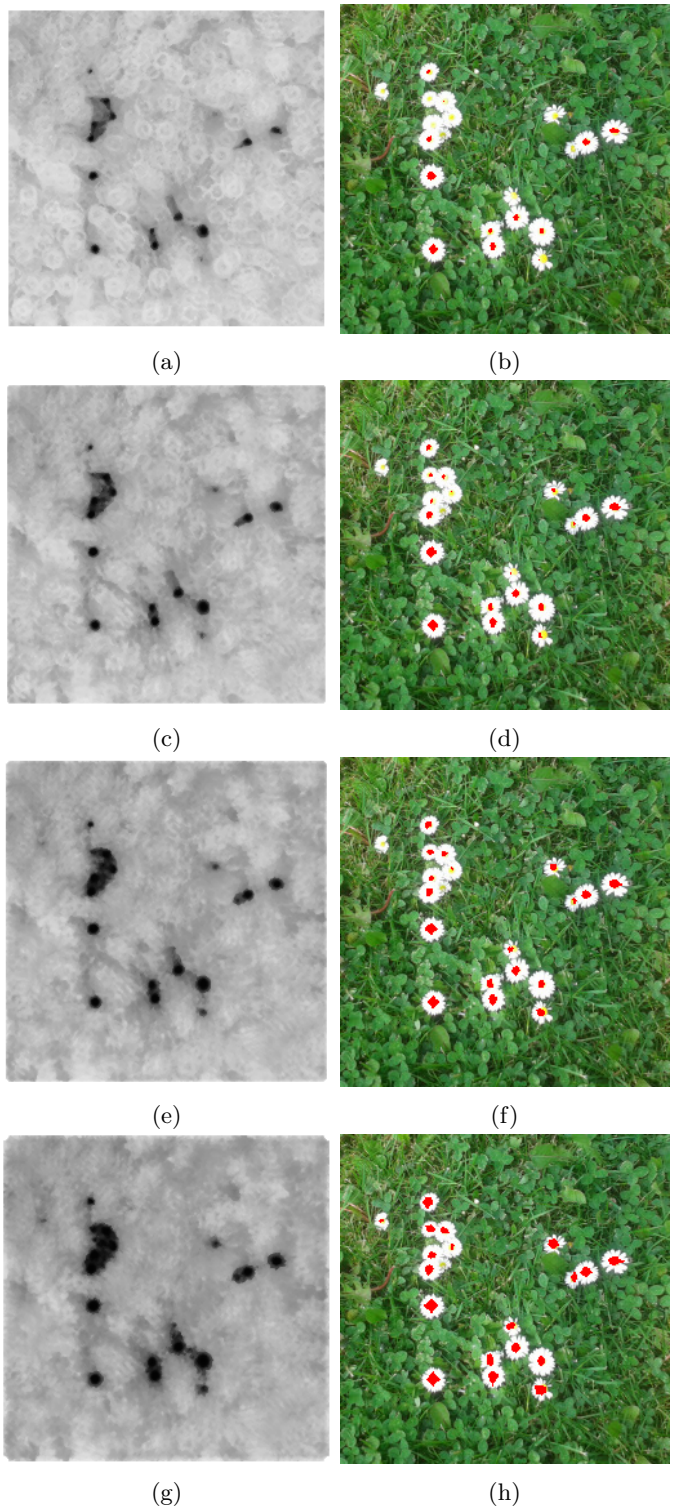


Figure 4.28: Foreground distance image and MDPOHMT result using h-similarity with **a,b)** $p = 90\%$, **c,d)** $p = 80\%$, **e,f)** $p = 70\%$, **g,h)** $p = 60\%$.

4.5.4 Application of the MDPOHMT to High Resolution Aerial Imagery

The MDPOHMT is designed to generalise to objects of interest over multiple images. In order to validate this, it was applied to a subset of aerial images from the DOTA dataset [269], Appendix A.1, to test its performance detecting various sizes and shapes of colour objects, in this case swimming pools. This class of object was chosen for the challenge that it presented with varying size, shape, colour, and levels of occlusion. In order to detect these objects, a generalised SE was designed through examination of examples and from prior knowledge of the objects of interest.

In order to generalise well to objects, including those which are noisy or partially occluded, the fit of an HMT often needs to be relaxed. This is preferable to creating multiple bespoke SEs as this not only takes time, but increases the time required to probe an image with these multiple SEs. Figure 4.29 shows the approach for estimating the appropriate PO parameters, using a tile of image P0059 from the DOTA dataset (Figure A.1.1).

Figure 4.29a shows the tile used to test the performance of the MDPOHMT. Figure 4.29b shows the colour SE used to probe this image and consists of a 7×7 square $\boldsymbol{S}_{FG}$ and a 51×51 hollow square $\boldsymbol{S}_{BG}$, with thickness 1. The result of performing the MDPOHMT on Figure 4.29a using the SE in Figure 4.29b with $p = 100\%$ and $T = 1$ is shown in Figure 4.29c.

A subset of pixels in each of the swimming pools in Figure 4.29a are successfully identified, however, the shapes of each of the swimming pools in the image are not reflected in the MDPOHMT output. Three pixels, each corresponding to a point in one of three swimming pools in the image, were selected to analyse using PO plots, and are located at each of the coloured crosses shown in Figure 4.29d. These selected points are close to the edges of each of the swimming pool objects as this represents the desired boundary of the object of interest. By analysing the PO plot in Figure 4.29e, an appropriate value for the PO parameter to detect each of the selected points can be set as $p = 90\%$, as denoted by the cross in Figure 4.29e.

The result of performing the MDPOHMT on Figure 4.29a where $p = 90\%$ is shown in Figure 4.29f, where a number of FPs have been introduced. By investigating the occupancy graph, shown in Figure 4.29g with a focussed view of the top left corner in Figure 4.29h, a foreground threshold value of $T = 0.2$ is applied, based on the distance between the image and $\boldsymbol{S}_{FG}$ at $p = 90\%$ when centred at these points. This has the effect of removing the FPs, as shown in Figure 4.29i, but also results in the pixel denoted by the blue cross in Figure 4.29d being excluded from the set of positively identified pixels. This is due to it having a foreground

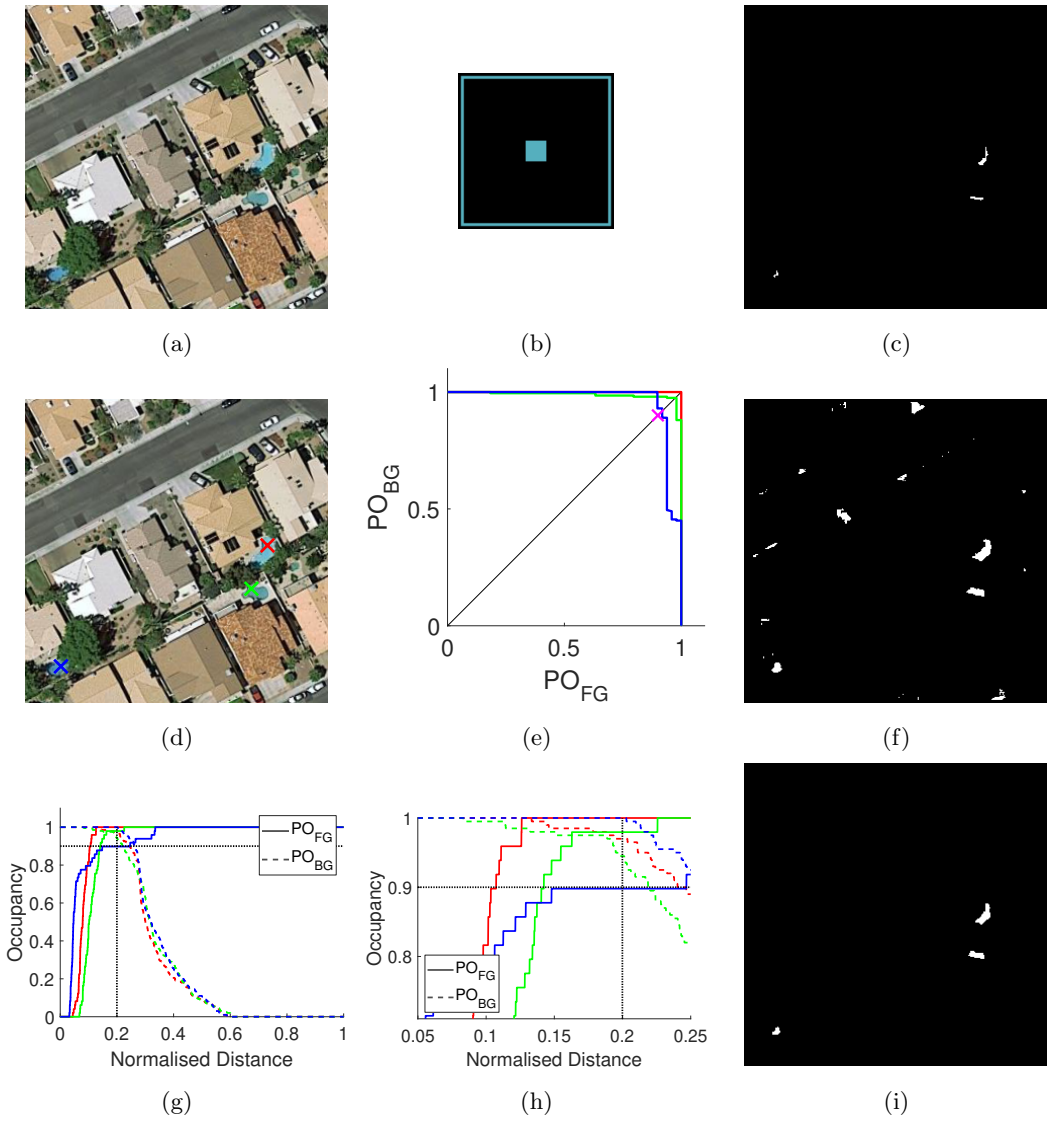
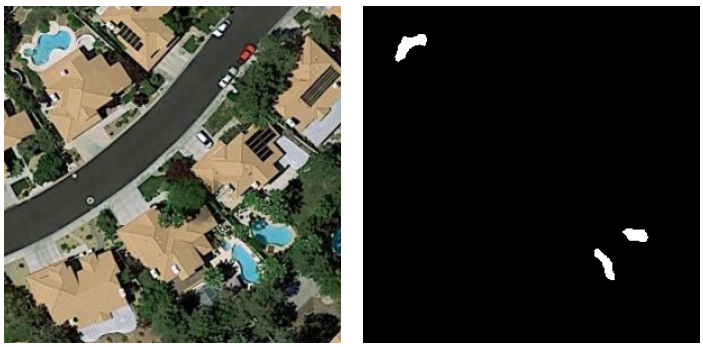


Figure 4.29: **a)** Test image. **b)** Composite SE used to detect the swimming pools in a). **c)** MDPOHMT result with no relaxation. **d)** Selected pixels containing swimming pools. **e)** PO Plots at the three selected pixels in d). **f)** MDPOHMT result with $p = 90\%$. **g)** Occupancy graphs at the three selected pixels in d). **h)** Zoomed version of g). **i)** MDPOHMT result with $p = 80\%$ and $T = 0.2$.

distance greater than that of the threshold at an occupancy value of $p = 90\%$, as $d_{FG} \approx 0.25$. Increasing T to account for this results in the reintroduction of false positives, as does reducing p without also reducing T further.

In addition to PO plot analysis, the optimal values of p and T can be refined by performing a grid search with multiple values for each parameter, $p = \{100\%,\ 95\%, \ldots\ 60\%,\ 55\%\}$ and $T = \{0.05,\ 0.075,\ 0.1\ \ldots\ 0.35,\ 0.375,\ 0.4\}$. By maximising the F1 score over multiple test

image tiles, optimal values for p and T can be found. For the purposes of this grid search and the results from Table 4.2, a TP refers to a successful MDPOHMT detection within the bounding box of each swimming pool in the DOTA dataset. It was found, through a mixture of this search method and PO analysis, that the optimal parameters for use in the MDPOHMT were $p = 80\%$ and $T = 0.15$ in the subset of images tested. Applying the MDPOHMT using these parameters to another test image, Figure 4.30a, provides the result shown in Figure 4.30b.

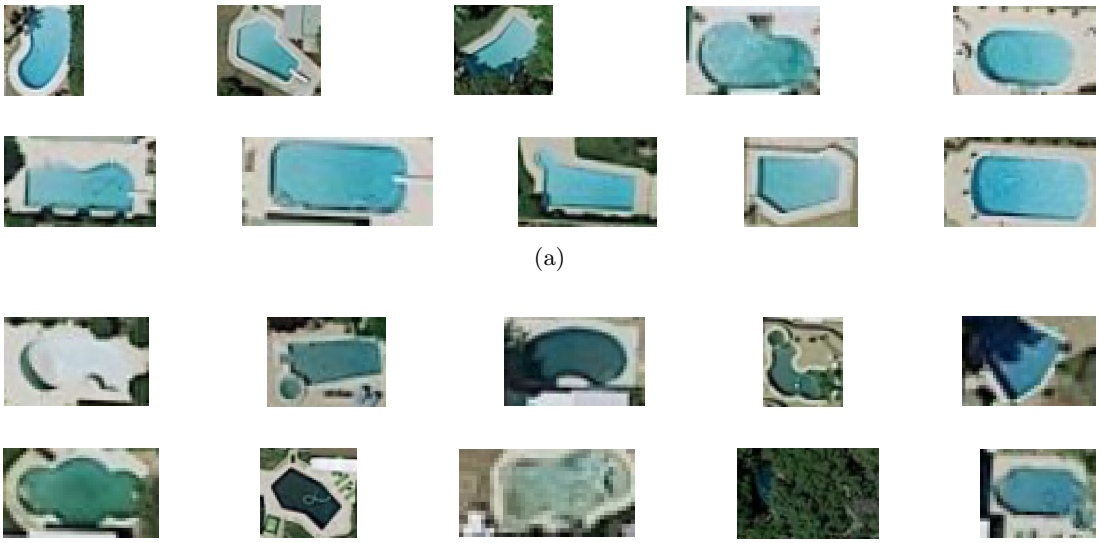


(a) (b)

Figure 4.30: **a** Sub-image from the DOTA dataset containing swimming pools. **b)** MDPOHMT result on a) with $p = 80\%$ and $T = 0.15$.

Using these parameters, the MDPOHMT can be applied to each of the DOTA images used in this thesis. Figure 4.31 shows the results of taking the MDPOHMT on a full image from the dataset, where detected swimming pools are outlined in green and missed pools in red. Falsely detected pixels are outlined in orange. Examples of the swimming pools that are detected and missed are shown in Figures 4.32a and 4.32b respectively. The full results from applying the MDPOHMT to the images from DOTA are shown in Table 4.2, where the image shown in Figure 4.31 is "*Image 7*".

Figure 4.31: Image from the DOTA test set used to validate the detection of swimming pools where objects detected using the MDPOHMT are highlighted in green and those missed are shown in red.



(a)

(b)

Figure 4.32: Examples of: **a)** positively detected swimming pools, **b)** missed swimming pools from Figure 4.31

Table 4.2: MDPOHMT and SHMT results on a subset of the DOTA Dataset using both single and multiple SEs.

MDPOHMT - $P = 80\%$, $T = 0.15$								
Image Number	1	2	3	4	5	6	7	Average
Processing Time (s)	4.2	2.5	1.1	0.9	2.8	2.1	2.6	-
No. True Positives	236	125	95	91	197	164	157	-
No. False Negatives	59	15	23	20	28	22	37	-
No. False Positives	0	1	0	3	2	0	5	-
Accuracy	80.00	89.29	80.51	81.98	87.56	88.17	80.93	84.06
F1 Score	0.8889	0.9398	0.8920	0.8878	0.9290	0.9371	0.8820	0.9081
Precision	1.0000	0.9921	1.0000	0.9681	0.9899	1.0000	0.9691	0.9885
Recall	0.8000	0.8929	0.8051	0.8198	0.8756	0.8817	0.8093	0.8406
MDPOHMT Multiple SEs - $P = 80\%$, $T = 0.15$								
Image Number	1	2	3	4	5	6	7	Average
Processing Time (s)	16.5	10.1	4.5	3.7	11.4	8.3	10.6	-
No. True Positives	264	132	102	96	202	173	168	-
No. False Negatives	31	8	16	15	23	13	26	-
No. False Positives	9	5	21	22	12	20	26	-
Accuracy	89.49	94.29	86.44	86.49	89.78	93.01	86.60	89.44
F1 Score	0.9296	0.9521	0.8465	0.8384	0.9203	0.9129	0.8660	0.8952
Precision	0.9670	0.9635	0.8293	0.8136	0.9439	0.8964	0.8660	0.8971
Recall	0.8949	0.9429	0.8644	0.8649	0.8978	0.9301	0.8660	0.8944
SHMT$_{0.9}$								
Image Number	1	2	3	4	5	6	7	Average
Processing Time (s)	22	14.2	6.4	5.3	15.9	11.7	14.6	-
No. True Positives	76	54	47	45	94	102	94	-
No. False Negatives	219	86	71	66	131	84	100	-
No. False Positives	0	0	0	0	0	0	0	-
Accuracy	25.76	38.57	39.83	40.54	41.78	54.84	48.45	41.40
F1 Score	0.4097	0.5567	0.5697	0.5769	0.5893	0.7083	0.6528	0.5805
Precision	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Recall	0.2576	0.3857	0.3983	0.4054	0.4178	0.5484	0.4845	0.4140
SHMT$_{0.9}$ Multiple SEs								
Image Number	1	2	3	4	5	6	7	Average
Processing Time (s)	88.6	56.4	25.5	20.6	62.7	46.8	58.3	-
No. True Positives	110	76	58	59	118	121	115	-
No. False Negatives	185	64	60	52	107	65	79	-
No. False Positives	2	2	0	1	2	1	5	-
Accuracy	37.29	54.29	49.15	53.15	52.44	65.05	59.28	52.95
F1 Score	0.5405	0.6972	0.6591	0.6901	0.6841	0.7857	0.7325	0.6842
Precision	0.9821	0.9744	1.0000	0.9833	0.9833	0.9918	0.9583	0.9819
Recall	0.3729	0.5429	0.4915	0.5315	0.5244	0.6505	0.5928	0.5295

On average, the method performs well and achieves an accuracy of 84% with an average F1 score of 0.9, precision of 0.98, and recall of 0.84. The number of FP detections is very low, with only eleven FPs in the set of seven images. The majority of objects are successfully detected, as indicated by the high number of TPs as well as the high accuracy and recall. Those that are not detected are often missed due to significant occlusion, shadow, or other variations in colour. The occlusions can be overcome by reducing the PO value further, however, this increases the risk of FPs if the transform is relaxed excessively. Alternative SEs can be designed by looking at examples that a single SE has missed in order to improve the detection rate and recall of the MDPOHMT. In this particular example, another three SEs were used to detect objects the first had missed, using information from the examples of FNs to design appropriate SEs. This process comes at the expense of execution time, highlighted by the results in Table 4.2. This has the desired effect of increasing the recall of the MDPOHMT from 0.84 to 0.89 at the expense of some precision (0.98 to 0.89) due to FPs introduced by the additional SEs, increasing from 11 to 115.

In order to compare the MDPOHMT with other related techniques, the SHMT was also applied to the same subset of the DOTA dataset with identical SEs to those used in the MDPOHMT. The results of the SHMT with a single SE and with multiple SEs are shown in Table 4.2. These results highlight the failure of the SHMT to generalise due to the need for the SE to fit the query image more exactly than the MDPOHMT, given the lack of spatial relaxation provided by using rank order filters. However, similarly to the results of the MDPOHMT, the addition of other SEs improves the results but the average F1 score is ~ 0.2 less than that achieved when using the MDPOHMT. This is highlighted in Figure 4.33 where the PR curves of the MDPOHMT, with $p = \{100\%,\ 80\%\}$ and $T = \{0.05,\ 0.075, ..., 0.375,\ 0.4\}$, and the SHMT, where $\eta = \{1,\ 0.95, ..., 0.55,\ 0.5\}$, using both single and multiple SEs are plotted.

Figure 4.33 shows the results of the MDPOHMT with both a PO value of $p = 100\%$ as well as a relaxed transform with $p = 80\%$. Both the MDPOHMT with $p = 100\%$ and SHMT perform similarly with a single SE, achieving high precision before decreasing as both T and η are relaxed. Using multiple SEs with the MDPOHMT improves the recall slightly, as indicated by the curve shifting to the right before tailing off. Using multiple SEs with the SHMT has a similar effect, with the recall being slightly increased, however, this is at the expense of precision with FPs introduced in both methods.

Examples of the FP detections introduced can be seen in Figure 4.34 and are mostly grey roofs with other blue objects within the image, these can ultimately be filtered out through further optimisation of the SEs. By reducing the PO value of the MDPOHMT to

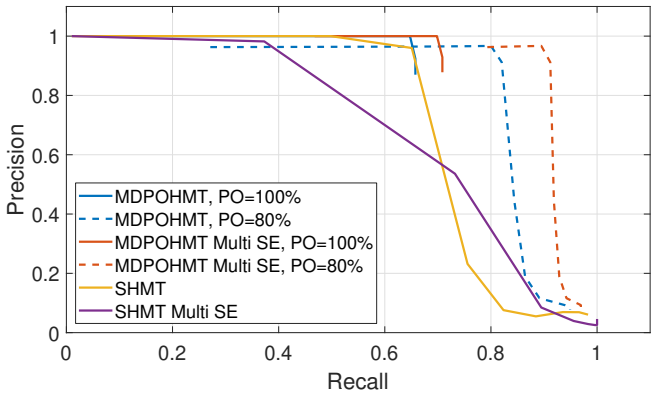


Figure 4.33: Precision-Recall curves of the SHMT with $\eta = 0.9$ and MDPOHMT where $p = \{100\% , 80\%\}$ with both single and multiple SEs.

80% (Figure 4.33), the recall was improved without a significant decrease in precision. This is true for both the single SE and multiple SE tests. From the results in both Table 4.2 and Figure 4.33 it is shown that the MDPOHMT offers an advantage over similar techniques. Table 4.2 additionally shows that the MDPOHMT has a faster execution time compared to the SHMT as the h-similarity measure is more computationally expensive to determine than a Euclidean distance as used in the MDPOHMT.

Figure 4.34: Examples of False Positive swimming pool detections when using multiple SEs with the MDPOHMT on the image shown in Figure 4.31.

4.5.5 Application to Hyperspectral Images

In this section, the ability of the MDPOHMT to generalise to N channel images is shown, firstly on synthetic images of varying complexity as well as on natural images from remote sensing platforms. It is also shown how the MDPOHMT can be applied more efficiently by applying dimensionality reduction to both the image and any probing SEs in order to perform target detection in an optimal subspace. This also aids in reducing the impact of Hughes phenomenon, *i.e.,* "*the curse of dimensionality*", by making each channel more informative.

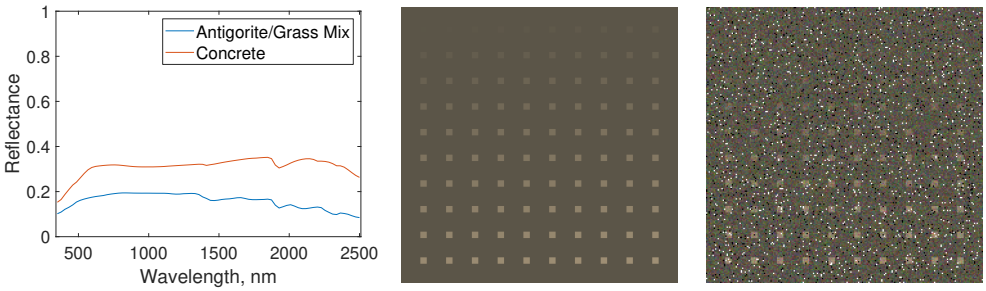
4.5.5.1 Synthetic Data

Another limitation of the SHMT, as discussed in Section 2.2.2, is that the model for similarity can only consider three channel colour images in either the RGB or L*a*b* colour spaces [40]. A general method for images in higher dimensionalities such as those present in MSI or HSI applications would be of benefit in object and target detection tasks. The reduced ordering method employed in the MDPOHMT is not restricted to $\mathbb{R}^3$, such is the case for RGB or other colour images, and can be readily applied to $\mathbb{R}^N$ images with no additional complexity, other than the additional dimensionality of the input data. While the Euclidean distance measure is used here in order to present the MDPOHMT, other distance measures, or indeed target detection methods such as the SAM or SID algorithms, Equations (3.56) and (3.61) respectively, which are explicitly suited to HSI applications, may be incorporated as the reduced ordering method within the MDPOHMT. Figure 4.35 shows a pair of synthetic images created in order to test the MDPOHMT on hyperspectral data.

Each of the images has a spatial dimension of 215×215 with 80 spectral bands and contain two Endmember spectra from the United States Geological Survey (USGS) Spectral Library [271], shown in Figure 4.35a, one (AMX25) comprised of a mixture of antigorite and grass which create the background and another consisting of concrete (GDS375) which forms the foreground at varying abundances throughout the image. The abundance measure was varied from 100% to 1%, as seen in Figure 4.35b in order to verify whether the MDPOHMT could be used in scenarios where there was no pure target spectra present in a pixel. In order to also validate the noise robustness of the MDPOHMT when applied to hyperspectral imagery, both Bipolar Impulsive noise with $\mathrm{P}(n) = 5\%$ as well as zero-mean Gaussian noise with $\sigma^2 = 1e^{-3}$ were added to the image (Figure 4.35c).

The MDPOHMT was first applied to the image that had not been perforated by noise (Figure 4.35b) with SEs designed using the desired target spectra from Figure 4.35a. The results of this operation are shown in Figure 4.36a where all 100 of the desired target objects are correctly detected with no FPs. Even those with little ($\leq 10\%$) abundance are detected as $\boldsymbol{S}_{FG}$ has marginally lower dissimilarity with these mixed regions than $\boldsymbol{S}_{BG}$ has with the background.

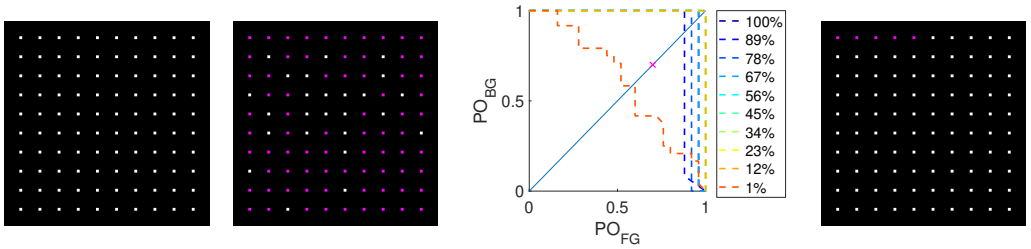
The MDPOHMT was then applied with $p = 100\%$ to the noise compromised image (Figure 4.35c) with the number of correctly identified objects dropping from 100 to 25, as seen in Figure 4.36b where white pixels indicate a TP detection and magenta pixels indicate a FN detection. This drop in recall can mostly be attributed to the impulsive noise added to the image as this can "puncture" the target regions of the image, causing the MDPOHMT to



(a) (b) (c)

Figure 4.35: Synthetic Images generated to assess the MDPOHMT for use on hyperspectral images. **a)** The two Endmember spectra used to create the image - the antigorite/grass mixed spectra forming the background and the concrete forming the in-painted targets. **b)** False-colour synthetic image created with 100 targets ranging from 100% abundance to 1%. **c)** False-colour synthetic image from b) with 5% Bipolar Impulsive noise and Gaussian noise with $\sigma^2 = 1e^{-3}$ and $\mu = 0$.

fail. Additionally, the Gaussian noise can cause the background to become more similar to the foreground and degrade the quality of the target spectrum. In order to rectify this, a number of example points were selected from the image with varying target abundances being represented. PO plots for each of these points were generated and are shown in Figure 4.36c. Using the PO plots at each pixel it was found that the 1% abundance target could not be detected due to the required PO falling below 50%. The other test objects could be detected with a PO value of $p = 70\%$ as represented by the magenta cross in Figure 4.36c. The results of the MDPOHMT in Figures 4.36a, 4.36b and 4.36d have been enlarged in order to increase their legibility, each detection is a single pixel.



(a) (b) (c) (d)

Figure 4.36: MDPOHMT results on the generated synthetic hyperspectral images **a)** MDPOHMT results on the image from Figure 4.35b with $p = 100\%$. **b)** MDPOHMT results on the image from Figure 4.35c with $p = 100\%$. **c)** PO plots at selected test pixels with decreasing target spectra abundance. **d)** MDPOHMT with $p = 70\%$ selected using the PO plot in c).

With this new occupancy value of $p = 70\%$, the MDPOHMT was reapplied to the noisy image. Upon relaxing the PO value, the number of positively detected objects rose from 25

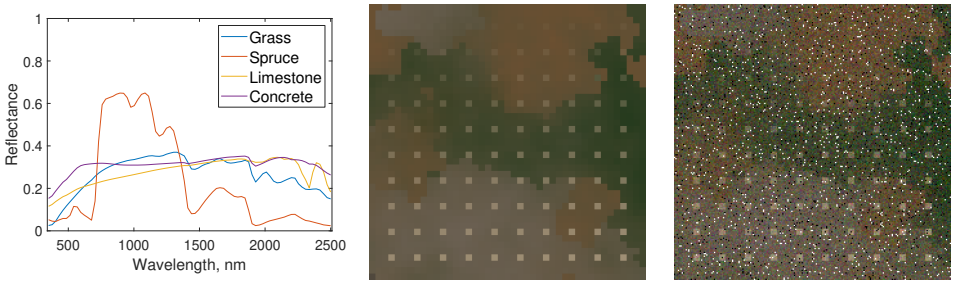
to 95, largely overcoming the noise present in the image. Only five objects were not detected when the noise was accounted for, each of those with spectral abundance $\leq 5\%$. This can be attributed primarily to the Gaussian noise overpowering the target spectra and blending it with the background spectra due to the low abundances of target spectra and subsequent high sensitivity to changes.

The spectra present in Figure 4.35a are near-uniform over the portion of the spectrum considered here. Thus, as seen in Figure 4.36a, they are relatively easy to delineate and detect with even the 1% abundance target being detected in the noise free image. Despite the Impulsive and Gaussian noise applied to the image, this uniformity is unrealistic and is intended to be an example of the extension of the MDPOHMT for application to spectral data. In order to assess how reliably the MDPOHMT can detect targets in a more complex environment, an additional pair of images was created using the USGS Spectral Library [271] along with the Hyperspectral Imagery Synthesis (EIAs) MATLAB toolbox [270] which allowed for the creation of randomised abundance maps for hyperspectral image synthesis.

Four spectra from the USGS Spectral Library were selected, three natural materials: dry grass (GDS480), spruce needles (ES-Needls-1), and limestone (CU02-11A) form the background of the synthetic scene and the same concrete spectrum from Figure 4.35 forms the in-painted targets with abundances decreasing from 100% to 1%. Each of the Endmember spectra present in the image are displayed in Figure 4.37a with both the grass and limestone spectra being similar to that of the target concrete. The background scene was generated using the Gaussian Fields method [270] and the targets were subsequently inserted, with an abundance of 100%, i.e., pure spectra, in the bottom right corner of Figure 4.37b and an abundance of 1% in the top left corner. The generated image is shown in Figure 4.37b and Figure 4.37c shows the image with additional Gaussian and Bipolar Impulsive noise generated using the same parameters as those in Figure 4.35c.

The MDPOHMT was first applied to the noise-free image shown in Figure 4.37b, the results of which can be seen in Figure 4.38a where 89 of the 100 targets are detected successfully despite the likeness between the targets and their background in places. When compared with the results on the first, simpler, synthetic image (Figure 4.35b), FPs are introduced. These FPs are represented as green pixels in the MDPOHMT results images. In total, 46 objects are detected incorrectly. One method for removing such FP detections is to lower the foreground distance threshold, T, however, this may also lead to a reduction in TP detections due to some targets having low abundance in the image and therefore less similarity to the probing SE.

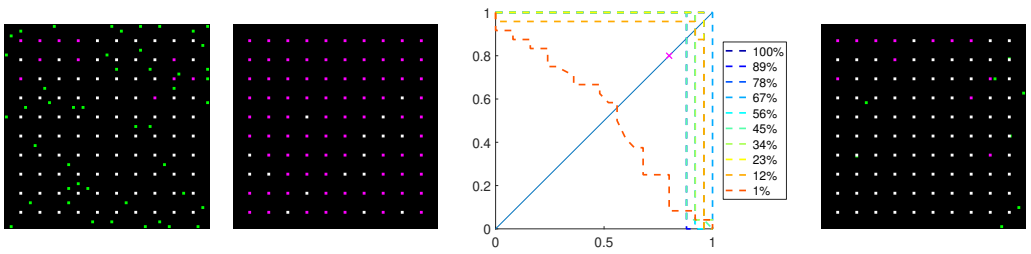
Applying the MDPOHMT with $p = 100\%$ to the noise degraded image (Figure 4.37c) yields



(a) (b) (c)

Figure 4.37: Hyperspectral image synthesis with a more complex, non-uniform, scene. **a)** The four Endmember spectra used to create the synthetic image. **b)** False-colour synthetic image created with 100 targets ranging from 100% abundance to 1%. **c)** False-colour synthetic image from b) with 5% Bipolar Impulsive noise and Gaussian noise with $\sigma^2 = 1e^{-3}$ and $\mu = 0$.

similar results to that of Figure 4.36b with only 16 of the 100 targets being detected, as shown in Figure 4.38b. This can be improved upon by lowering the PO, with the appropriate value set using a PO plot (Figure 4.38c). Through investigation, a PO value of $p = 80\%$ (represented by the magenta cross in Figure 4.38c) was sufficient to detect each of the targets tested, excluding the target with 1% abundance which, once again, required a PO value of $p < 50\%$. Applying the MDPOHMT to the noisy image with $p = 80\%$ results in the image shown in Figure 4.38d where 88 of the 100 targets are detected with only 12 FN and 8 FP detections. By relaxing the percentage fit through reducing the PO value based on the PO plot analysis, the Bipolar Impulsive noise is better accounted for.



(a) (b) (c) (d)

Figure 4.38: **a)** MDPOHMT results on the image from Figure 4.37b with $p = 100\%$. **b)** MDPOHMT results on the image from Figure 4.37c with $p = 100\%$. **c)** PO plots at selected test pixels with decreasing target spectra abundance. **d)** MDPOHMT with $p = 80\%$ selected using the PO plot in c).

4.5.5.2 Performing the MDPOHMT in a Subspace

In order to reduce the computational requirements and processing time when applying the MDPOHMT to hyperspectral data, the transform can be performed in an optimal

dimensionality reduced subspace. As the MDPOHMT, and HMTs in general, require SEs in order to detect objects of interest, the SEs have to be projected into the same reduced domain as the image. As discussed in Section 3.5.1, PCA [77] and its variants offer a method of projecting auxiliary, or supplementary, variables which were not present during the calculation of the PCs into the same reduced domain using a coefficient or projection matrix, given in Equation (3.40). This allows for the image and associated SEs to be projected into the same reduced domain by the image's coefficient matrix, $\boldsymbol{V}_{PCA}^{-1}$, calculated using Equation (3.39), for processing as show in Equations (4.12) and (4.13):

$$\hat{\boldsymbol{I}} = \boldsymbol{V}_{PCA}^{-1}\boldsymbol{I} \tag{4.12}$$

$$\hat{\boldsymbol{S}} = \boldsymbol{V}_{PCA}^{-1}\boldsymbol{S} \tag{4.13}$$

Where $\boldsymbol{I}$ and $\boldsymbol{S}$ are the image and SE respectively and $\hat{\boldsymbol{I}}$ and $\hat{\boldsymbol{S}}$ are their projections. While PCA provides an optimal set of compression coefficients, it is not the most efficient method as it requires the eigenvalues and eigenvectors to be calculated on a per-image basis. This is similar for each of the DR methods from Section 3.5 so some future investigation may find a set of coefficients for a set of images that was representative enough to only be calculated for one, or multiple, image(s) from the set. By projecting each image and the corresponding set of SEs using one set of representative coefficients, a more efficient process can be defined. The current proposed method is shown graphically in Figure 4.39 while a flowchart of the proposed methods is shown in Figure 4.40.

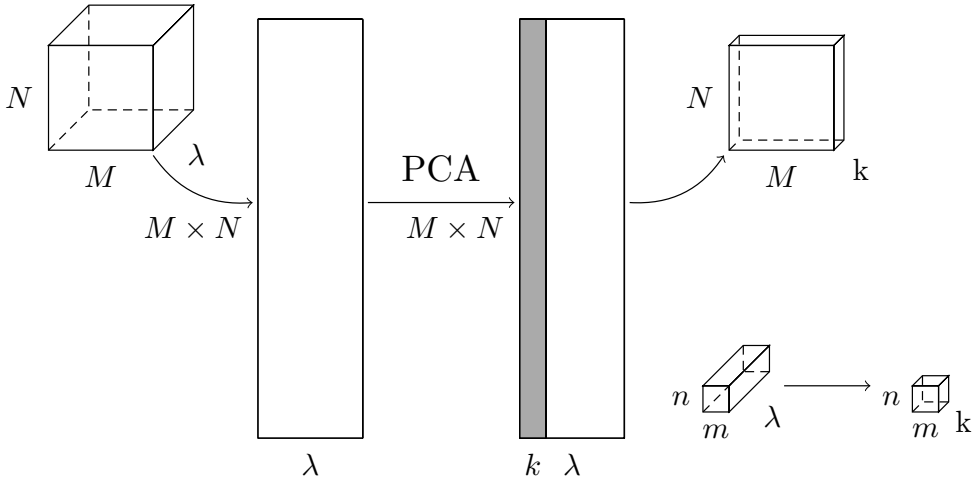


Figure 4.39: Proposed dimensionality reduction scheme using PCA to project an image and any relevant SEs into the same reduced-dimension domain.

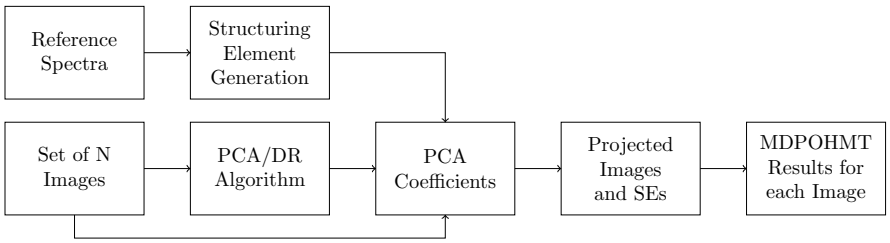


Figure 4.40: Flowchart of the proposed method for spectral dimensionality reduction for a set of images and their SEs.

PCA and its variants, along with many other forms of DR seek to exploit spectral redundancy. With large hyperspectral images this is desirable to overcome the so called curse of dimensionality. However, spatial redundancy is often exhibited in remote sensing applications and target/anomaly detection, wherein much of the image is vegetative and of little interest or importance to an object/target detection task. By using measures such as NDVI, pixels containing large amounts vegetation can identified. NDVI is a simple and effective technique in determining the amount of vegetation present in each pixel. The intensities at two wavelengths are used to obtain a ratio that determines the NDVI, one in the red part of the spectrum, λ_{Red}, where light is absorbed by vegetation and another in the NIR, λ_{NIR}, where it is reflected. These wavelengths and the formula for NDVI and two of its variants are listed in Table 3.1.

The effects of adding DR to the MDPOHMT were tested using the images from the OP7 dataset supplied by BAE systems (Figure A.3.1). Each of these images show the same scene containing three atmospheric calibration panels, used as the targets for this example. These images are discussed further in Section 5.2. The spectra of each of these panels was known and appropriate SEs were designed. The spectra for each of the three targets in the OP7 datasets, as well as three background classes, are shown in Figure 4.41 along with the bands used for calculating NDVI.

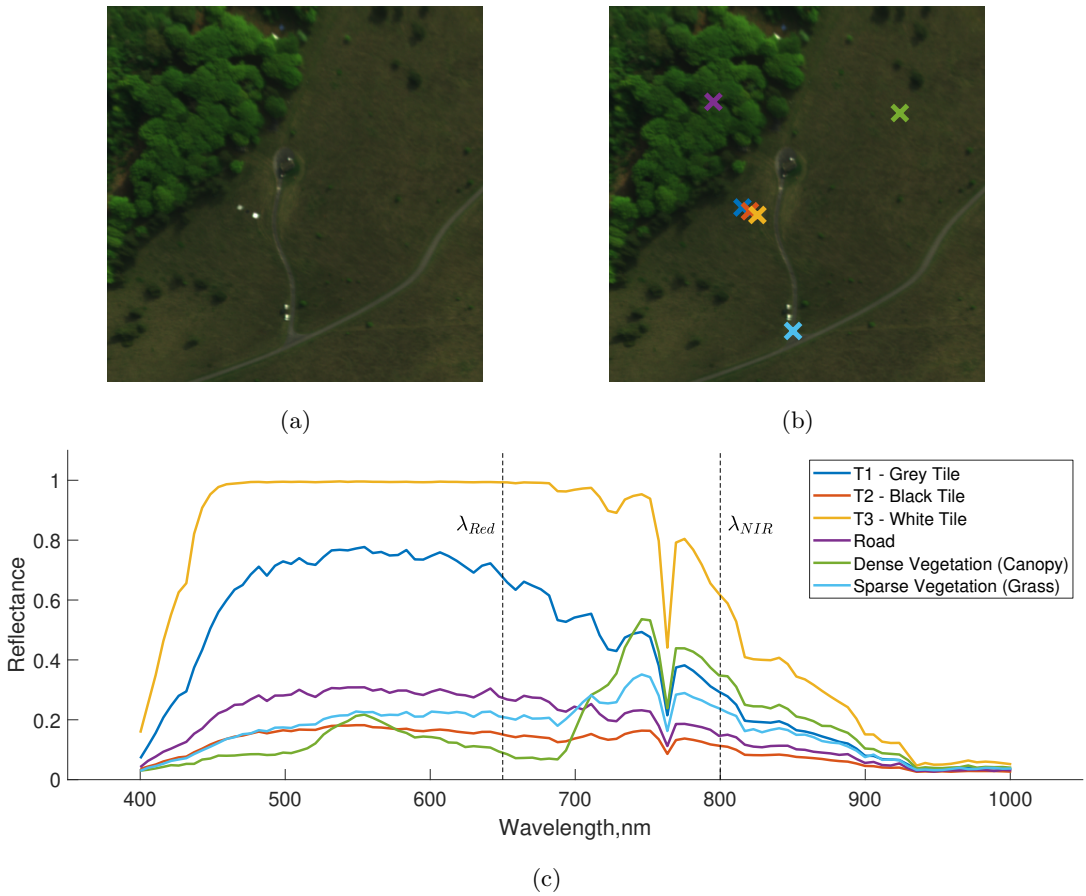
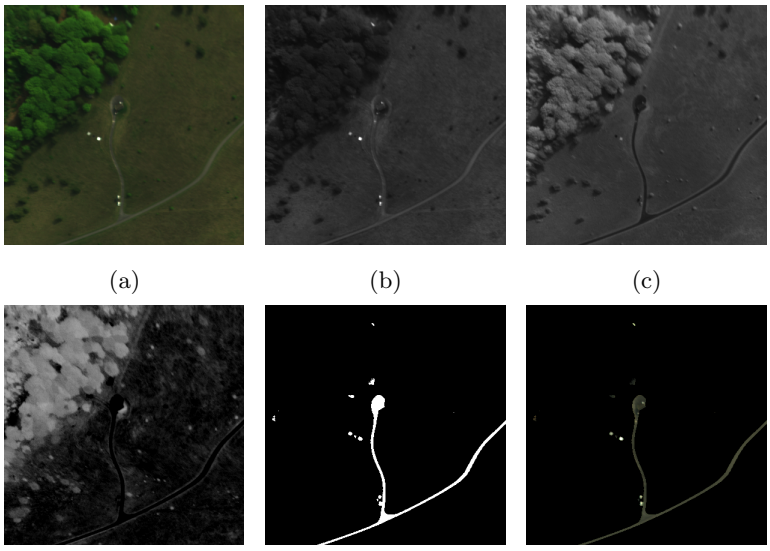


Figure 4.41: Examples of vegetative and non-vegetative spectra and NDVI. **a)** Image from the OP7 dataset. **b)** Locations of the spectra detailed in c). **c)** Set of vegetative and non-vegetative spectra and bands used in calculating NDVI.

From Figure 4.41, it is possibly to see how NDVI can discriminate between vegetative and non-vegetative pixels. The vegetative pixels sampled each have an increased reflectance value in the NIR region of the spectrum when compared to the red, whereas the synthetic targets and road pixels are either equally or more absorbent.

The values of NDVI can be thresholded and used to mask out regions of interest in the image, reducing the number of PUTs to be queried using the MDPOHMT. The stages of this NDVI segmentation are shown in Figure 4.42, where Figure 4.42a shows the pseudocolour representation of the image and Figures 4.42b and 4.42c show the red and NIR bands of the image respectively. Brighter pixels in the resultant NDVI image indicate a higher likelihood that the pixel contains vegetation, as can be seen in Figure 4.42d. These NDVI values can be thresholded to isolate only non-vegetative pixels, as shown in Figure 4.42e, and used to mask the original image, as in Figure 4.42f.

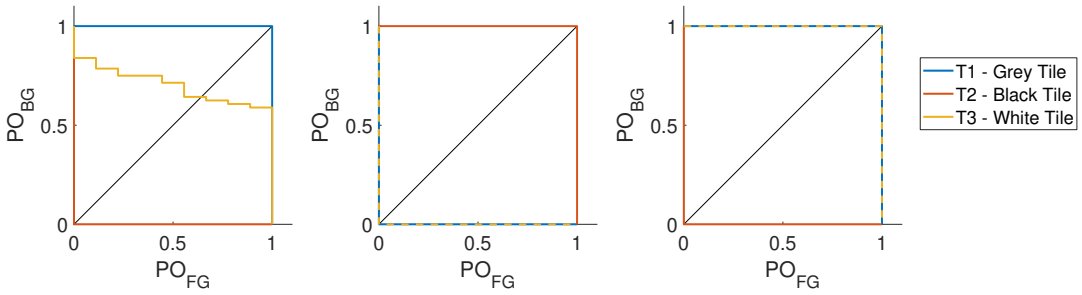


(a) (b) (c)

(d) (e) (f)

Figure 4.42: Stages of NDVI segmentation: **a)** Pseudo-colour image of the OP7 image scene **b)** Image at λ_{Red} **c)** Image at λ_{NIR} **d)** NDVI measurement of the scene. **e)** Binary mask for ROI **f)** Masked pseudo-colour image

In order to apply the MDPOHMT to the image from Figure 4.41a, using the spectra corresponding to each target from Figure 4.41c, a composite SE for each target object was created. In each case, this consisted of a 3×3 square $\boldsymbol{S}_{FG}$ and a 9×9 hollow square $\boldsymbol{S}_{BG}$ with thickness of 1 pixel. This allowed for PO plots at each of the target locations using each composite SE, $\boldsymbol{S}$, to be generated, and they are displayed in Figure 4.43.

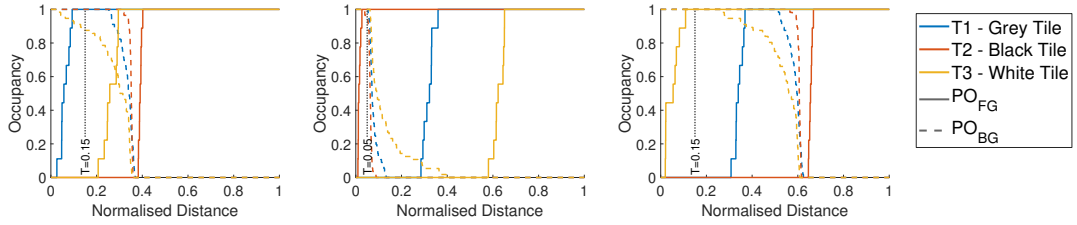


(a) (b) (c)

Figure 4.43: PO plots of image 1 from the OP7 dataset where each composite SE is translated to the location of **a)** target 1 "grey tile", **b)** target 2 "black tile", and **c)** target 3 "white tile".

From Figures 4.43a and 4.43b, when translating each of the designed SEs to the centre of targets 1 and 2, only the SE designed to detect these targets produce a positive detection at $p = 100\%$. However, when the SEs designed to detect target 1 is translated to the location of

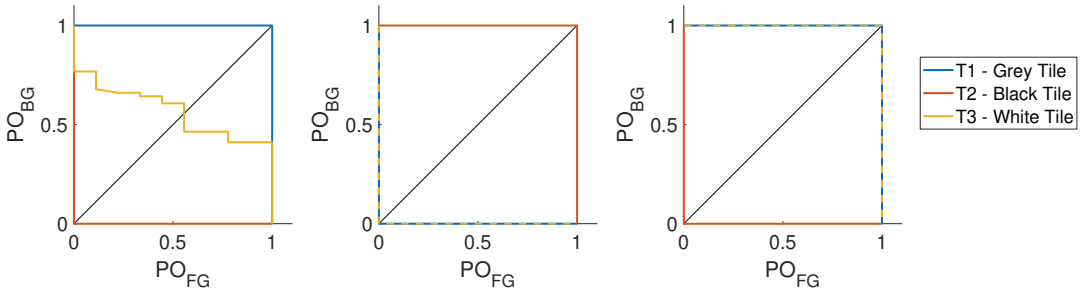
target 3, the MDPOHMT returns a positive detection with $p = 100\%$, as shown in Figure 4.43c. This occurs as the grey $\boldsymbol{S}_{FG}$ is closer to the spectrum of the white target than the corresponding $\boldsymbol{S}_{BG}$ is to the spectrum of the background. Similarly to the examples displaying the detection of traffic lights in RGB imagery from Figure 4.18, this can be corrected by setting a suitable value for the foreground distance threshold, T, using the corresponding occupancy plots. These are shown in Figure 4.44 where the threshold value applied to each SE is denoted as a vertical black line. Despite the correct detection when using the SEs designed for targets 1 and 2, a threshold is used in the MDPOHMT for each of the probing SEs to suppress FP detections.



(a) (b) (c)

Figure 4.44: Occupancy plots of image 1 from the OP7 dataset where each composite SE is translated to the location of **a)** target 1 "grey tile", **b)** target 2 "black tile", and **c)** target 3 "white tile".

In addition to applying the MDPOHMT on the raw data, the query image and associated SEs can be projected into the PC domain using Equations (4.12) and (4.13) respectively. PO plots can be generated in the PC domain in an identical manner to those generated in the raw full dimensionality space. The PO plots when translating each of the projected SEs to each of the target locations in the projected image are shown in Figure 4.45.

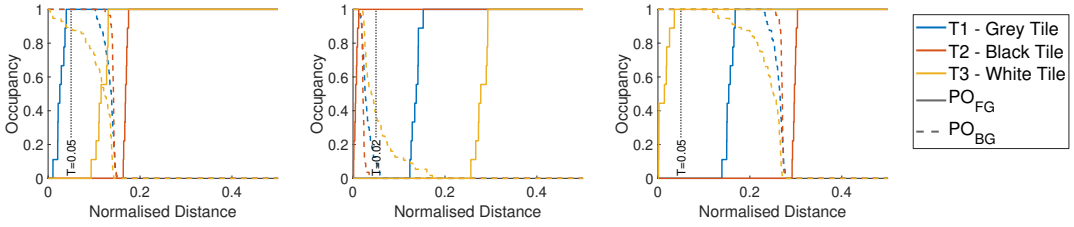


(a) (b) (c)

Figure 4.45: PO plots of image 1 from the OP7 dataset in the PC domain where each composite SE is translated to the location of **a)** target 1 "grey tile", **b)** target 2 "black tile", and **c)** target 3 "white tile".

The PO plots in the PC domain indicate the same problem as those in the original high dimensionality space, targets 1 and 2 are able to be detected solely by the the SEs designed

for them, whereas target 3 is detected by both the SEs for target 3 and target 1, as shown in fig. 4.45c. This can again be corrected by setting the foreground threshold, T, appropriately using the occupancy plots shown in Figure 4.46.



(a) (b) (c)

Figure 4.46: Occupancy plots of image 1 from the OP7 dataset in the PC domain where each composite SE is translated to the location of **a)** target 1 "grey tile", **b)** target 2 "black tile", and **c)** target 3 "white tile".

The parameters derived from each of the PO and occupancy plots for each target in both the full dimensionality and PC spaces are given in Table 4.3. The results of applying the MDPOHMT on each of the raw and projected images in the OP7 dataset, both with and without NDVI masking, are shown in Figure 4.47.

The ground truth locations of the targets in each image are shown in Figures 4.47a to 4.47c, where the colours of each marker correspond to the spectra from Figure 4.41 and the PO and occupancy plots from in Figures 4.43 to 4.46. Figures 4.47d to 4.47f show the results of performing the MDPOHMT using each of the designed SEs on the full dimensionality image with the relevant parameters from Table 4.3. Figures 4.47g to 4.47i show the results when projecting each image and SE into the optimal PC domain for processing. In both cases, targets 1 and 3 are able to be detected without any associated FPs. However, due to the dark spectrum exhibited by target 2, where little spectral information is available, there are a number of FPs introduced. By reducing the number of locations to perform the MDPOHMT, this not only decreases the execution time, but filters out any unwanted locations based on known spectral characteristics. Figures 4.47j to 4.47l show the result of performing the MDPOHMT on the full dimensionality images, with NDVI determining which pixels to query, essentially acting as a mask on the detections from Figures 4.47d to 4.47f. Similarly, Figures 4.47m to 4.47o show the results of the MDPOHMT on the PC projected image with the NDVI-based masking. The

Table 4.3: MDPOHMT parameters used to detect targets in the OP7 dataset.

	Target 1	Target 2	Target 3
Image space	$p = 100\%$, $T = 0.15$	$p = 100\%$, $T = 0.05$	$p = 100\%$, $T = 0.15$
PC domain	$p = 100\%$, $T = 0.05$	$p = 100\%$, $T = 0.02$	$p = 100\%$, $T = 0.05$

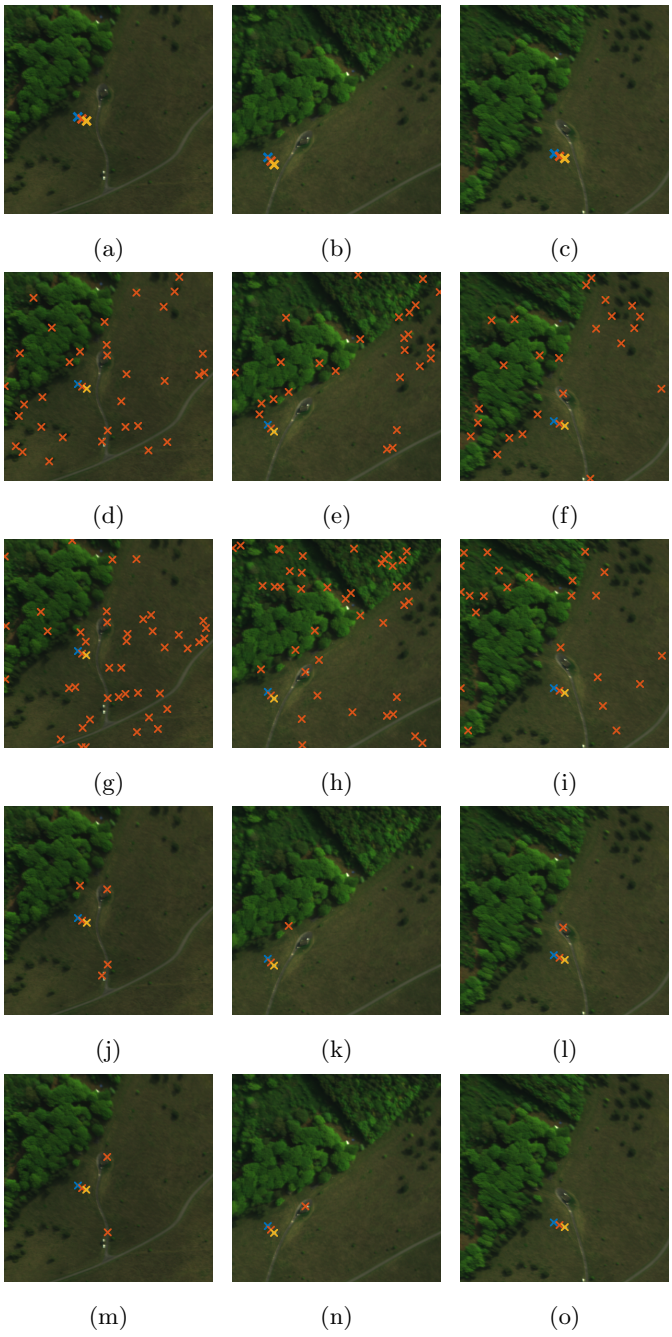


Figure 4.47: Results of the dimensionality reduction on the MDPOHMT **a-c)** Ground truth locations. **d-f)** Results of the raw MDPOHMT. **g-i)** Results of the MDPOHMT with PCA applied. **j-l)** Results of the MDPOHMT with NDVI applied. **m-o)** Results of the MDPOHMT with both PCA and NDVI applied.

effect of the PCA-based DR on the F1 scores obtained both with and without NDVI acting as a subset of pixels to sample are given in Table 4.4.

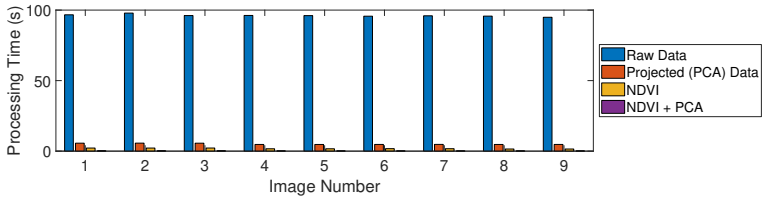
Table 4.4: F1 Score achieved using the MDPOHMT with each DR scheme on the OP7 dataset.

	Pre-processing			
	Raw	PCA	NDVI	PCA + NDVI
OP7 1	0.73	0.72	0.91	0.95
OP7 2	0.75	0.72	0.97	0.98
OP7 3	0.75	0.74	0.96	0.98
Average	0.74	0.73	0.95	0.97

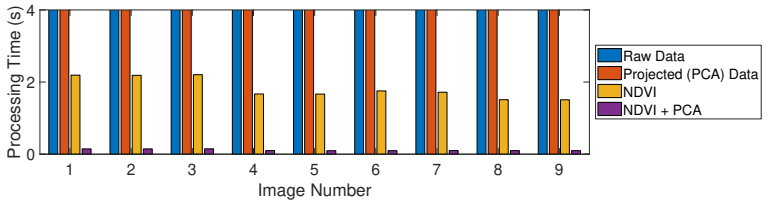
The precision of the MDPOHMT increases as the FPs in the original image are reduced. This reduction can be attributed to the refinement of the areas of interest using NDVI. Using an appropriate threshold for distance based on *a priori* knowledge of the object of interest or by *a posteriori* measurement, the MDPOHMT can be used as a target detection measure, both in the full dimensionality image domain or in the projected PC space. The resultant timings for each method are shown in Table 4.5 and in the bar chart in Figure 4.48.

Table 4.5: Comparison of execution time and effects of each DR method on the OP7 dataset.

	Pre-processing			
	Raw	PCA	NDVI	PCA + NDVI
Execution Time (s)	96.87	5.65	2.19	**0.14**
No. Pixels retained	16800000	640000	331170	**12616**
% of Total Pixels	100%	3.81%	1.97%	**0.075%**



(a)



(b)

Figure 4.48: **a)** Execution time of the MDPOHMT with the various DR techniques. **b)** Magnified version of a).

As expected, performing the MDPOHMT on a projected image using PCA-based DR greatly reduces the run time compared to that when applied to the raw image. As much of the computation time is allocated to querying every pixel in the image, the use of NDVI to remove vegetative pixels of little interest in a TD application further decreases the time taken to obtain an MDPOHMT result. Applying both NDVI and PCA further reduce the run-time as both the dimensionality of the data and the number of points the SE is translated to are reduced.

4.6 Summary

The extension of Mathematical Morphology to colour images presents a difficult and undefined challenge, the tools provided by morphology have been shown to be highly useful in many image processing applications and therefore the necessity for general extensible tools is of value.

In this chapter, a novel extension of the Morphological HMT, the Multi-Dimensional Percentage Occupancy Hit-or-Miss Transform, is formalised. The MDPOHMT is intended as a general technique which extends the morphological HMT for use in object detection and pattern recognition applications in colour and multivariate images. This is achieved using a reduced ordering scheme which reduces multivariate data into a scalar value via some distance or similarity measure. The proposed transform was validated, first using synthetic colour images in order to display the correct operation of the technique followed by object detection and discrimination experiments on natural images. The MDPOHMT was tested on the DOTA aerial image dataset in order to display how the technique performs in a large dataset with objects of interest varying in size, shape, colour, texture, and occlusion levels. The MDPOHMT was also compared with other similar greyscale and colour HMTs with the proposed method performing favourably.

In order to make the MDPOHMT more robust to noise, rank-order filters are used in place of the hard maximum and minimum filters commonly employed in morphological techniques. This allows for the transform to be flexible in cases where the image contains noise or where objects of interest are occluded. By relaxing the constraints of traditional erosion and dilation operations, objects with a high enough percentage fit to the corresponding probing SEs can be detected. By utilising a percentage fit, the transform remains unbiased to changes in size or shape of the objects of interest or the SEs. The parameters used in this relaxation are intuitive and can can be determined through trial and error using only a few training samples or indeed estimated and refined empirically, by using the extended PO plots introduced in this work. Using this notion of PO and rank order filters combined with the use of PO plots

for setting the appropriate parameters, the MDPOHMT has been shown to perform favourably when compared with related single channel and multivariate HMT definitions on the considered pattern recognition tasks.

Through experimentation using the DOTA dataset it was found that the MDPOHMT operator, in conjunction with the Euclidean distance-based reduced ordering, can be very sensitive to changes in colour. The Euclidean distance measure was used as an example of the reduced ordering necessary for the operation of the proposed technique. Other distance measures may be more appropriate for specific applications, especially in the case of HSI where dedicated spectral measures may be implemented. Similarly, alternative colour spaces, such as L*a*b* or HSV, may be more suited to colour morphology as discussed in the literature. A general way of overcoming this sensitivity to changes in colour or spectral information in objects of interest is to simply design multiple composite SEs, one for each image or object. However, this can be an involved and bespoke process, reducing the generalisation of the technique. By using multivariate SEs and encoding them with the desired colour information, as well as size and shape, the issue of colour perception is limited - as it simply needs to match the SE - while also allowing for non-flat objects to be detected with a single composite SE. As with the vast majority of morphological operators the SE choice and design is still an active area of research and optimal design of SEs for particular applications is of interest in future work. An alternative way to relax the sensitivity to changes in object colour could be to have a multivariate threshold or decision space rather than a scalar distance threshold. When testing the MDPOHMT, it was found that, like other HMT definitions, the MDPOHMT itself is not scale invariant. This is something that may be suited for future investigation but can be overcome by simply creating multiple SEs based on the scale and orientation of the objects to be detected, or by designing SEs which can generalise between objects.

While in this chapter, the focus is solely on the extension of the HMT in particular, the methods used here may be extended and applied to other morphological operators such as erosions and dilations and, subsequently, openings and closings. Similarly to the techniques described in Section 3.8, the MDPOHMT may be integrated into a machine learning framework. This may have the effect of augmenting feature extraction either as an individual HMT layer or alternatively using the notion of PO and rank-order filters within the feature pooling stage.

Chapter 5

Spatial and Spectral Dimensionality Reduction Schemes for Improved Target Detection

5.1 Introduction

Imagery from aerial and satellite based remote sensing platforms has become increasingly prevalent and a necessary source of information in various areas of research including determining land use [110], assessing vegetation health [229], and smart agriculture [273, 274], amongst others. In addition to these, it is used extensively in military and security applications such as disaster relief [275], target tracking [276] and classification, and change detection [128, 277]. Target Detection from aerial imagery is a major challenge and is an active area of research within the disciplines of signal and image processing [127, 278, 279]. There have been a wide range of TD algorithms of varying complexities developed over the last few decades [129]. The most notable difficulties in aerial TD are discussed in [121] and include sensor noise effects, and atmospheric attenuation and subsequent correction which can both lead to variabilities in target signature. These effects are noted, however in this thesis, it is assumed they are negated prior to any processing as such algorithms are beyond the scope of this research.

Depending on the system, remote sensing data can consist of high resolution RGB colour data, radar, multispectral images, or hyperspectral images. The latter, while providing a great deal of useful information, often at wavelengths beyond the range of human vision, introduces a vast quantity of data which must be handled and processed. Dimensionality Reduction techniques offer methods of compressing and remapping this high dimensionality data into a reduced, and sometimes more informative, uncorrelated subspace. As hyperspectral images contain high levels of redundancy they are easily compressed using sparsity-based approaches [280] or by applying DR methods, as discussed in Sections 2.3 and 3.5. High-dimensionality data is often sparsely distributed across some space and as a result, statistical meaning is

harder to derive or less generalisable – the so-called "curse of dimensionality". Reducing the dimensionality of the data in a meaningful way can reduce its sparsity and improve the distribution across some reduced dimensionality space, allowing for more informative analyses to be inferred. Coupling spectral DR with TD in order to improve detection and classification rates has been covered widely in the literature [84, 120–124], as well as in the author's personal research [125, 126] (Appendix C), and has been shown to improve the performance of TD and classification algorithms.

In TD applications, targets are often sparsely positioned in an imaged scene, therefore large amounts of spatial redundancy are exhibited as most of the spatial information in the scene is of little or no interest. This spatial redundancy, like the spectral redundancy also present in hyperspectral images, can be exploited in order to attain increased performance and efficiency. In the previous chapter (specifically Section 4.5.5.2), using NDVI to spatially mask hyperspectral images in order to constrain the region of interest in the scene was investigated. This spatial masking was applied on raw, full dimensionality, images as well as on compressed representations created with DR techniques. However, in this chapter it is refined such that the spatial DR is applied prior to the calculation of the dimensionality reduced image in order to removed redundant information from the calculation of the optimal projection and refine the subspace in which any TD is performed.

NDVI and its variants are most often used in remote sensing applications to quickly and effectively assess vegetation health and abundance [229] within an imaged scene. Here, their use in determining vegetation content on a pixel-wise basis was used to provide a measure of how informative a pixel may be or how likely it is to hold a target signature. Pixels are categorised as informative or non-informative with the non-informative pixels being discarded. By removing such pixels, the DR calculation was simplified by decreasing the number of observations, whilst also simplifying and suppressing the background class. As TD algorithms can be represented as a binary classification, improving the separation between target and background classes consequently improves TD performance [278].

The Joint Spatio-Spectral Dimensionality Reduced Target Detection pipeline proposed in this chapter is shown in Figure 5.1 and consists of three stages. First, spatial DR is applied. In this thesis, due to the scenes investigated being predominantly vegetative, VIs are used to mask out the background, reducing the number of pixels contributing to the further steps. Secondly, this reduced set of pixels are used to perform optimal spectral DR. By excluding the background, the remaining subset of pixels, containing the targets and non-vegetative background, will be better represented in the reduced subspace following DR-based projection. Lastly, having

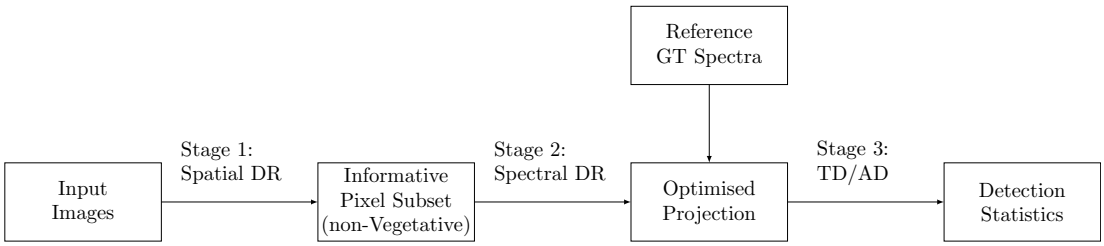


Figure 5.1: Proposed JSSDR-TD pipeline.

projected the target reference spectrum into the same domain, TD can be applied within the subspace to identify objects of interest.

In this chapter, the use of coupled spatial and spectral DR for enhancing hyperspectral TD applications is proposed and its effects on the accuracy of TD algorithms are investigated. This proposed JSSDR-TD approach aims to decrease both the spatial and spectral redundancy exhibited in hyperspectral images, improving the efficiency and performance of various benchmark TD algorithms. The proposed JSSDR-TD method is then tested on two hyperspectral datasets containing multiple targets in varied scenes. The remainder of this chapter is structured as follows; the justification for a JSSDR-TD approach is first discussed, followed by a selection and verification of the optimal spatial DR, spectral DR, and TD algorithms. Then, the results of applying the JSSDR-TD pipeline to real image data are shown and discussed.

5.2 Joint Spatial and Spectral Dimensionality Reduction for Improved Subspace Target Detection

HSI provides a wealth of information about a scene but, as discussed in Sections 2.4 and 3.5, the images gathered using HSI are often highly redundant and, as a result, highly compressible. This redundancy is exhibited both spectrally, where successive images are taken at neighbouring bands and therefore have high correlation, and spatially, where very little of the information across the scene is of interest in certain use cases. Examples of images used in this thesis which exhibit such spatial redundancy are shown in Figure 5.2.

(a) (b) (c) (d)

Figure 5.2: False colour images from the datasets used in this chapter. **a)** OP7 image. **b)** Target Region of a). **c)** UDRC Selene image. **d)** Target Region of c).

Figure 5.2 shows an example image from each of the OP7 (Figure A.3.1) and UDRC Selene (Figure A.3.2) datasets. These images both display a vegetative scene with multiple target objects with varying spectral signatures present. The OP7 image, shown in Figure 5.2a, consists of predominantly vegetation, with a path, structure, and a dense treeline also present. Three spectral calibration panels (grey, black, and white) are used as targets of interest, with the target region being highlighted with a red square and cropped in Figure 5.2b. There are also spectrally similar objects both in the treeline and below the target region. The UDRC Selene image, shown in Figure 5.2c, is similarly predominantly vegetative, however, there is a larger proportion of non-vegetative pixels and assorted clutter present in the scene compared to the images from the OP7 dataset. The target area is highlighted by a red square and cropped in Figure 5.2d. In these images, there are up to seven materials of interest as listed in Table 5.5 (pg. 178) along with their distribution across each of the UDRC Selene images used.

5.2.1 Dimensionality Reduction Validation

With each of the DR algorithms listed in Section 3.5, one key attribute to investigate is how well, when projected, the ground truth adheres to the observed measurement. As the data

is projected from the original domain to a reduced dimensionality subspace, subtle differences in reference and observed spectra may lead to larger disparities in the projected space. As the target is sparse within the image, coefficients which preserve its features may not be prioritised. Figure 5.2 shows the disparity between the ground truth measurement and an observed measurement, corresponding to the same target material, in the original image domain, as well as for each of the four DR schemes, in order to observe the effect of transforming auxiliary information into these reduced subspaces.

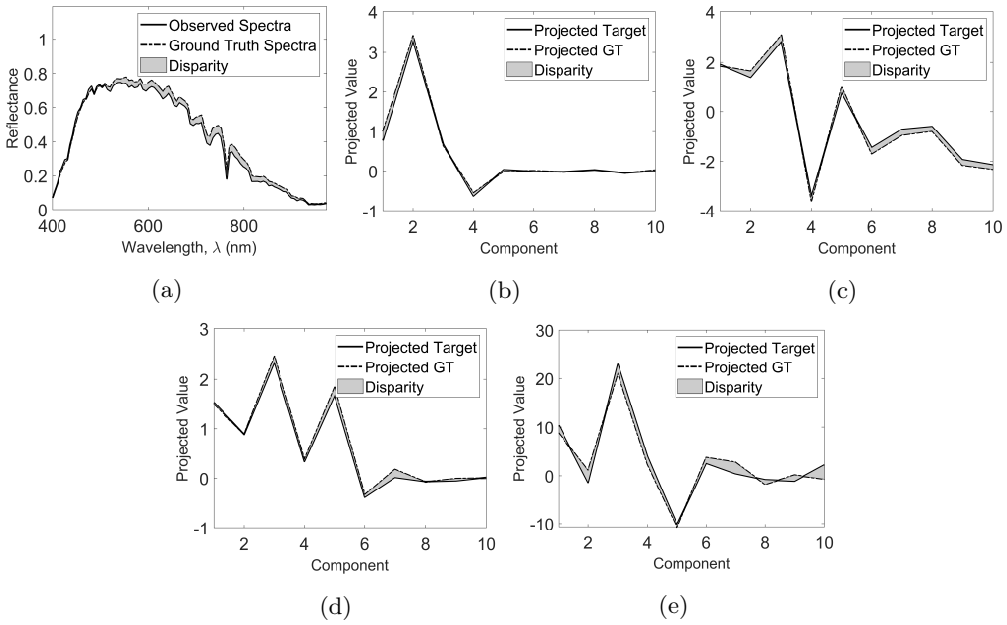


Figure 5.3: Comparison between the observed and ground truth measurements between each of the DR methods; **a)** raw, **b)** PCA, **c)** MNF, **d)** FPCA, and **e)** ICA.

Each of the DR methods tested ensure the observed spectrum adheres well to the ground truth spectrum. In order to quantify this, the area between the observed and reference spectra was calculated and normalised based on the range of values observed in each projected space. The resultant disparity between each of the observed spectra/projected measurements from Figure 5.3 are compared in Figure 5.4.

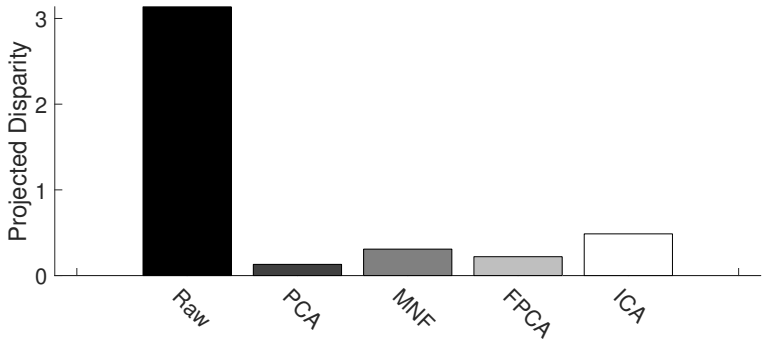


Figure 5.4: Disparity between the observed and ground truth spectra using DR

From Figure 5.4 it is shown that each of the DR methods perform better than the raw data when considering the disparity between the observed measurement and the ground truth in each domain. This is due to the raw data having ~ 10 times the components of each of the projected versions. The disparity between the projected ground truth and the projected observed spectrum is highest when using ICA, as shown in Figure 5.4, with PCA achieving the lowest disparity in the projected subspace.

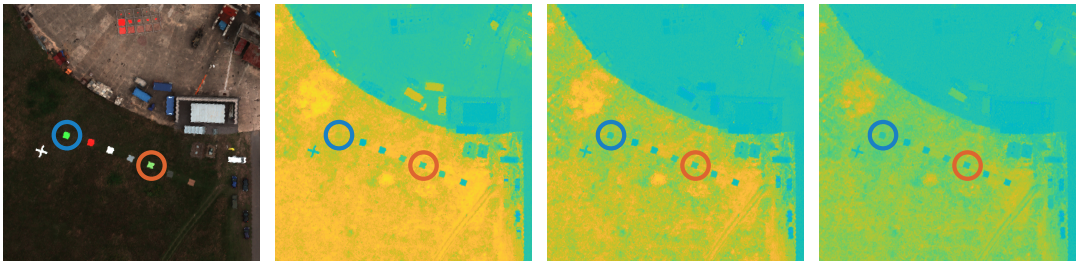
5.2.2 Spatial Dimensionality Reduction using Vegetation Indices

As discussed in Section 3.5.2, three VIs, each a variant of NDVI, are considered for use as spatial dimensionality reduction in this thesis. These are NDVI [228], NDVI_{re} [109, 110], and RENDVI [111, 112] and are presented in Table 3.1.

In order to assess which VI gave the best separation between vegetative and non-vegetative pixels, the ground truth spectra of multiple green targets from the Selene dataset as well as the average spectra of a patch of vegetation were investigated. Figure 5.5 shows the test image used as well as the results of each of the three VIs, with the locations of the ground truth materials, "green perspex" and "green ceramic" circled in blue and orange respectively in Figure 5.5.

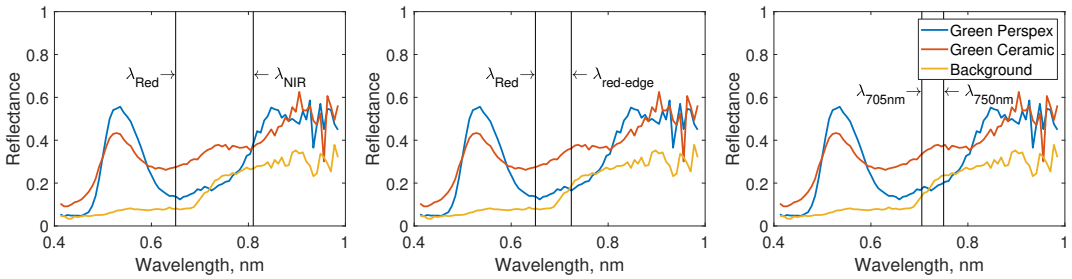
The ground truth spectra of the two target materials as well as the average background spectrum are shown in Figure 5.6. Each of the three VIs investigated produce a ratio between the intensity of a pixel at two bands. The wavelengths used to calculate each VI are visualised in Figures 5.6a to 5.6c for NDVI, NDVI_{re}, and RENDVI respectively, with the resultant VI values for the two targets, and the background spectrum, are shown in Table 5.1.

The distribution of VI values when applying each of the methods investigated to the image from Figure 5.5a are displayed in Figure 5.7 where clear bimodal distributions corresponding to non-vegetative and vegetative pixels are obtained. Additionally, a threshold for separating



(a) (b) (c) (d)

Figure 5.5: Comparison of VIs for spatial DR. **a)** UDRC Test Image. **b)** NDVI ratio. **c)** $NDVI_{re}$ ratio. **d)** RENDVI ratio. (In **b)** to **d)** warmer colours indicate higher levels of vegetation and colder colours indicate non-vegetation.)



(a) (b) (c)

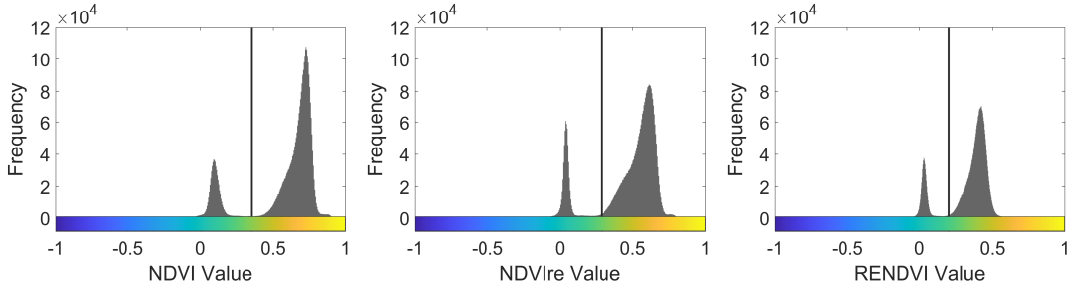
Figure 5.6: Wavelengths used to calculate each VI and three test spectra. **a)** NDVI wavelengths. **b)** $NDVI_{re}$ wavelengths. **c)** RENDVI wavelengths.

the modes of each distribution using Otsu's method [281] is shown as a vertical line in each histogram.

From Figures 5.5 and 5.6 and Table 5.1, it is possible to see that NDVI and RENDVI have lower separability between the "green perspex" target and the background. In fact, it can be observed that the green perspex target is near indistinguishable from the background in Figure 5.5b with only six of the seven targets having a low enough NDVI value to be separated from the background. The result of applying a binary mask, obtained via automatically thresholding each of the VI images using Otsu's method are shown in Figure 5.8. Despite having a distinct spectral profile, as shown in Figure 5.6a, the green perspex has an almost identical NDVI value to the background (0.48 vs 0.53) indicating the ratio between the two NDVI bands is nearly the same. By replacing the NIR band with one in the red-edge portion of the spectrum, as is the case when using $NDVI_{re}$, a much greater separation is achieved (0.09 vs 0.39). This is due to the red-edge phenomenon, when the intensity of the background spectra rises sharply reflecting NIR light, correlating well with the presence of healthy plant matter. RENDVI, while successfully segmenting all seven targets in this example, creates a

Table 5.1: Vegetation index ratios obtained for background and targets

Vegetation Index	Green Perspex Ratio	Green Ceramic Ratio	Background Ratio
NDVI	0.48	0.14	0.53
$NDVI_{re}$	0.09	0.13	0.39
RENDVI	0.10	0.06	0.28

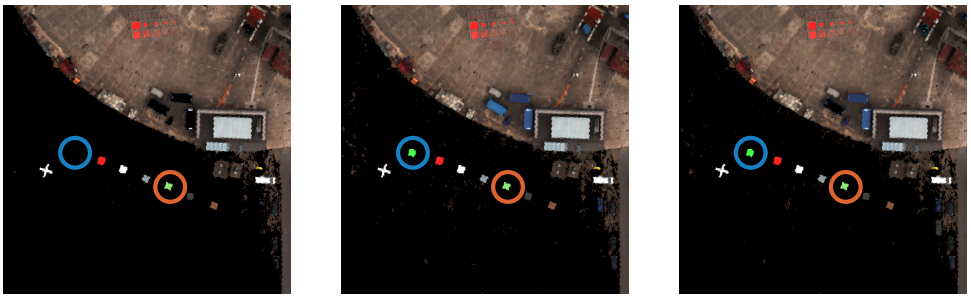


(a) (b) (c)

Figure 5.7: Histograms of each of the VI results with an automated threshold calculated using Otsu's method **a)** Histogram and threshold of NDVI. **b)** Histogram and threshold of $NDVI_{re}$. **c)** Histogram and threshold of RENDVI.

lower contrast between background and target when compared with $NDVI_{re}$.

The case for using $NDVI_{re}$ is strengthened when investigating the distributions of the three VIs intensities from Figure 5.7. All three VIs exhibit a bimodal-like distribution with peaks for the vegetative and non-vegetative pixels. RENDVI has the tightest spacing between the two modes whereas both NDVI and $NDVI_{re}$ can be more easily separated into vegetative and non-vegetative modes. The non-vegetative distribution when using $NDVI_{re}$ has a smaller standard deviation and is more homogeneous which allows it to be separated more effectively. As a result, $NDVI_{re}$ is used to implement a form of spatial DR in this chapter.



(a) (b) (c)

Figure 5.8: VI-based masking using automatic thresholding of **a)** NDVI, **b)** $NDVI_{re}$, and **c)** RENDVI.

5.3 Subspace Refinement using Spatial Dimensionality Reduction

The aim of any DR technique is to represent the original high dimensionality data in a lower dimensional subspace, reducing the complexity of the data whilst retaining, or extracting, the most important or informative features. DR algorithms will therefore attempt to represent every point in input data, regardless of its utility in subsequent processing. As a result, by excluding information of known unimportance (with regards to target detection) such as definite background pixels, the efficiency and efficacy of DR may be improved. In this chapter, spatial DR, in the form of NDVI_{re}, is used as a mask for selecting which pixels are included in the DR calculation. This is in contrast to how NDVI was applied as a post-processing step in Section 4.5.5.2 to filter out erroneous detections after PCA-based DR had been applied.

By removing these features and calculating the optimal representation of the remainder, the DR does not account for and represent unnecessary variance in the reduced dimension space. This has potential to improve the appearance of desirable features in the projected subspace by accentuating the variance between classes without the overhead of representing a, potentially overwhelming, background class. This also has the added benefit of reducing the processing time of the DR calculation as the number of observations is reduced.

An image from each of the OP7 and UDRC datasets was selected to test the use of spatial DR prior to spectral DR, both of these image datasets are discussed further in Section 5.5. From each of the images, 200 random points of non-vegetative, *i.e.,* residual background points not removed by the spatial DR stage of the JSSDR-TD pipeline, were selected alongside the target ground truth spectra to compare how spatial DR affects the calculation of the spectral DR. Residual, non-vegetative, points were selected such that there was a 1-to-1 comparison between the projected points with and without the subspace refinement step following spatial DR.

The PC plots given in Figure 5.9 show the results of performing PCA on the image from the OP7 dataset Figure A.3.1a with and without NDVI_{re}-based background removal shown in Figures 5.9c, 5.9e and 5.9g and Figures 5.9b, 5.9d and 5.9f respectively. Each point is coloured according to its membership of one of the background or target classes, and an ellipse characterising the 95% confidence interval of each cluster is drawn as a class boundary in this PC space.

Figures 5.10 and 5.12 show the first three PC images for each of the datasets projected using PCA with and without spatial DR. In both of the images tested from these datasets,

using spatial DR to refine the PCA calculation increases the separation of the desired target spectra in the PC domain – both from one another and the residual background not removed in the spatial DR step. This is most notable in Figure 5.9 with target 2 (orange), which is the most similar to the background, where by applying spatial DR prior to spectral DR, there is a slight separation in the new projected domain allowing for improved detection, as discussed in Section 5.5. Additionally, as visible in Figures 5.11e and 5.11g compared to Figures 5.11d and 5.11f, there is much increased separation from the background for each of the targets, with the exception of target 7, where only two of the PCs (Figure 5.11b) characterise the target spectra without utilising spatial DR. This is also visible in Figure 5.12e which highlights very little variation between the targets and background compared to Figure 5.12f where the individual targets continue to be characterised in the projected domain after background clutter removal.

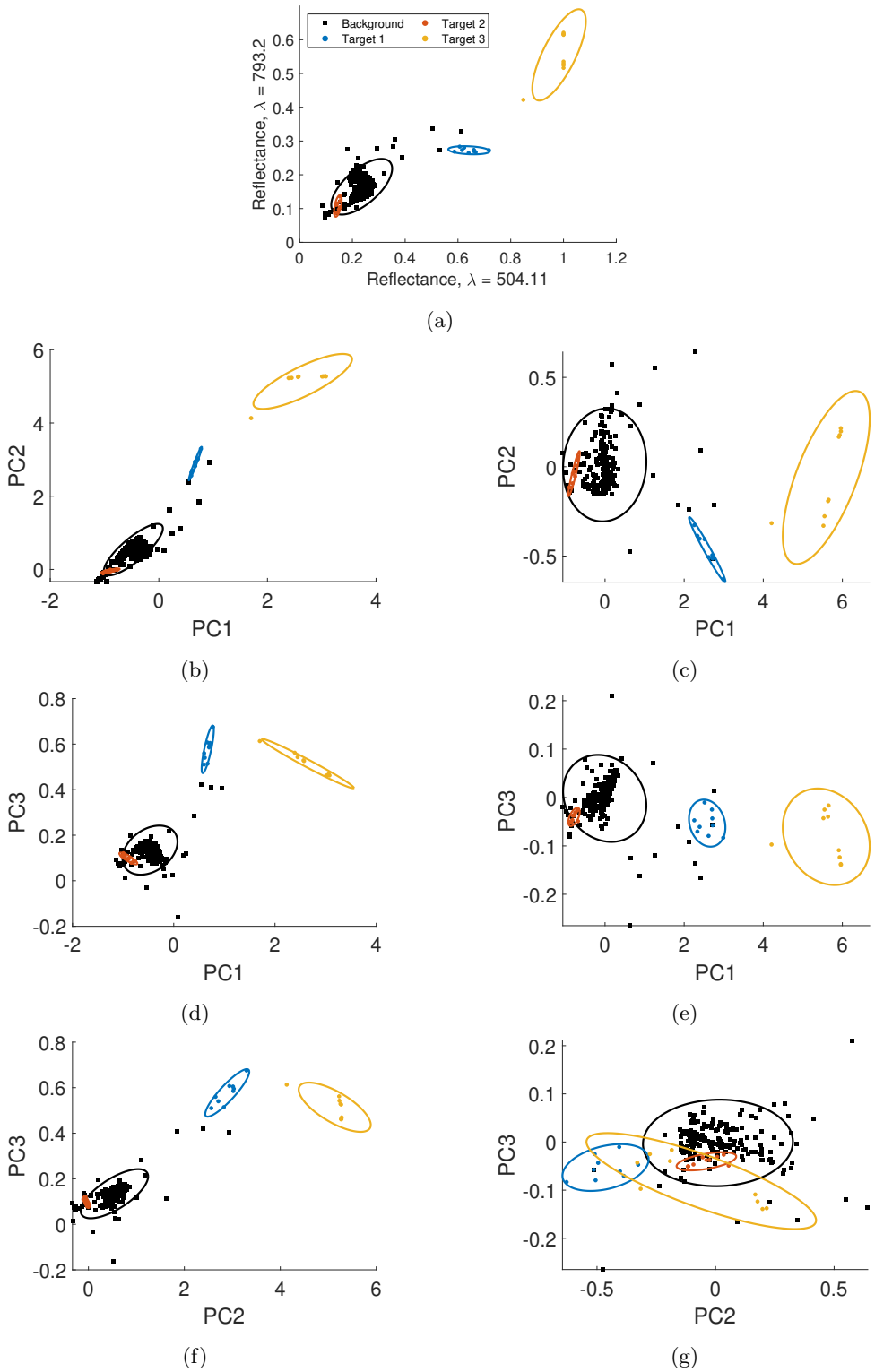


Figure 5.9: PC Plots of the OP7 Dataset excluding and including background removal. **a)** Query spectra in the original image domain, **b,d,f)** PC plots without background removal, **c,e,g)** PC plots with background removal.

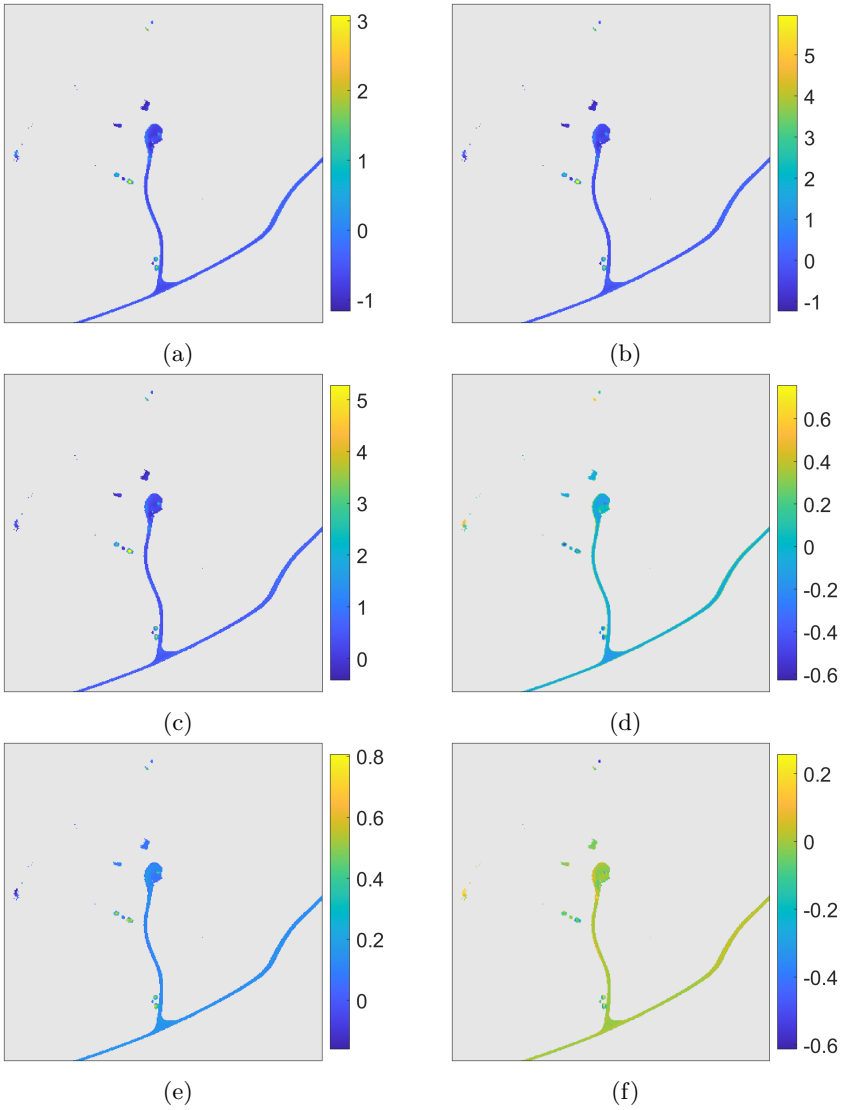


Figure 5.10: Principal Component Images of the OP7 Dataset. **a)** PC1 without background removal. **b)** PC1 with background removal. **c)** PC2 without background removal. **d)** PC2 with background removal. **e)** PC3 without background removal. **f)** PC3 with background removal.

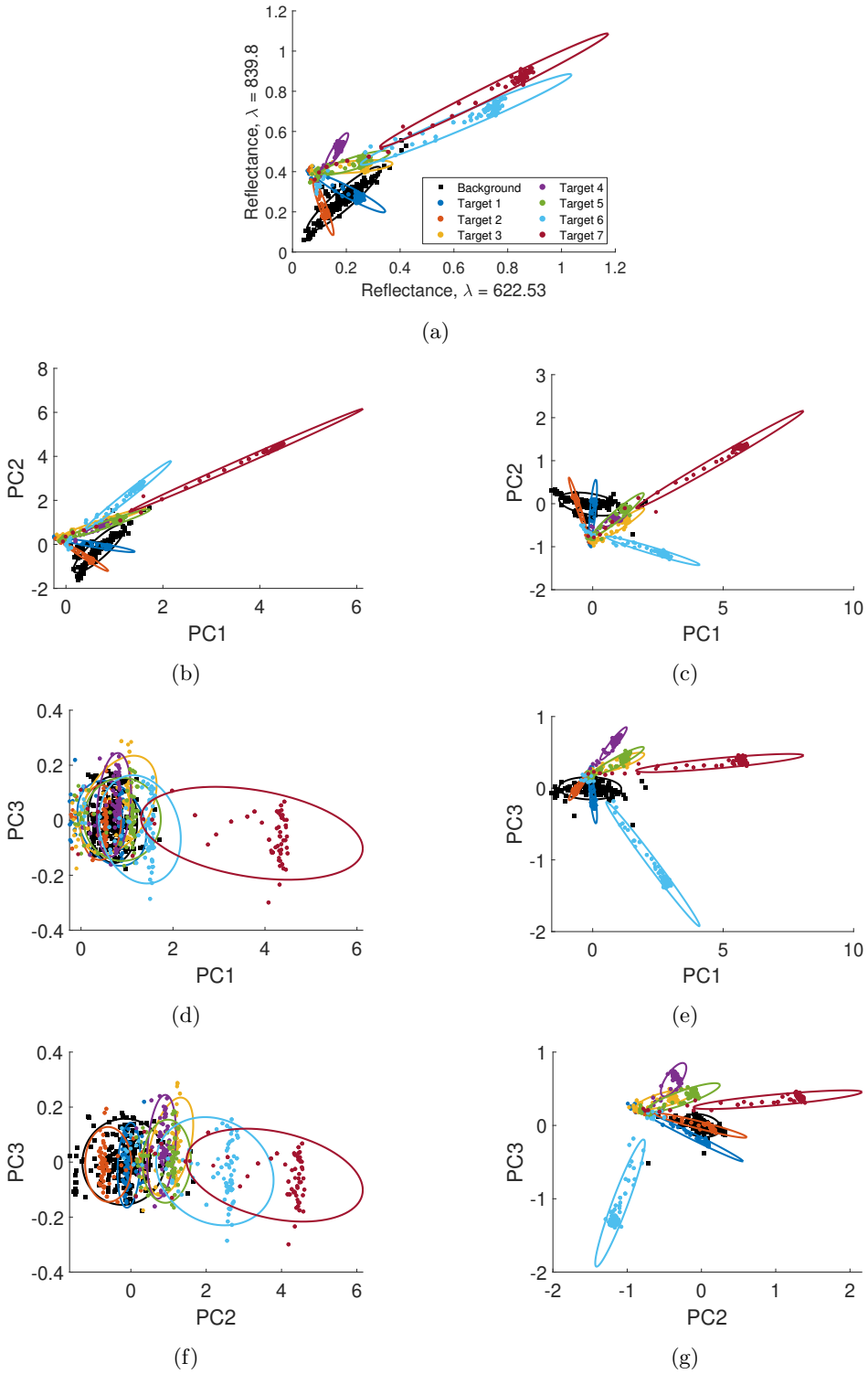


Figure 5.11: PC plots of an image from the UDRC dataset excluding and including background removal. **a)** Query spectra in the original image domain, **b,d,f)** PC plots without background removal, **c,e,g)** PC plots with background removal.

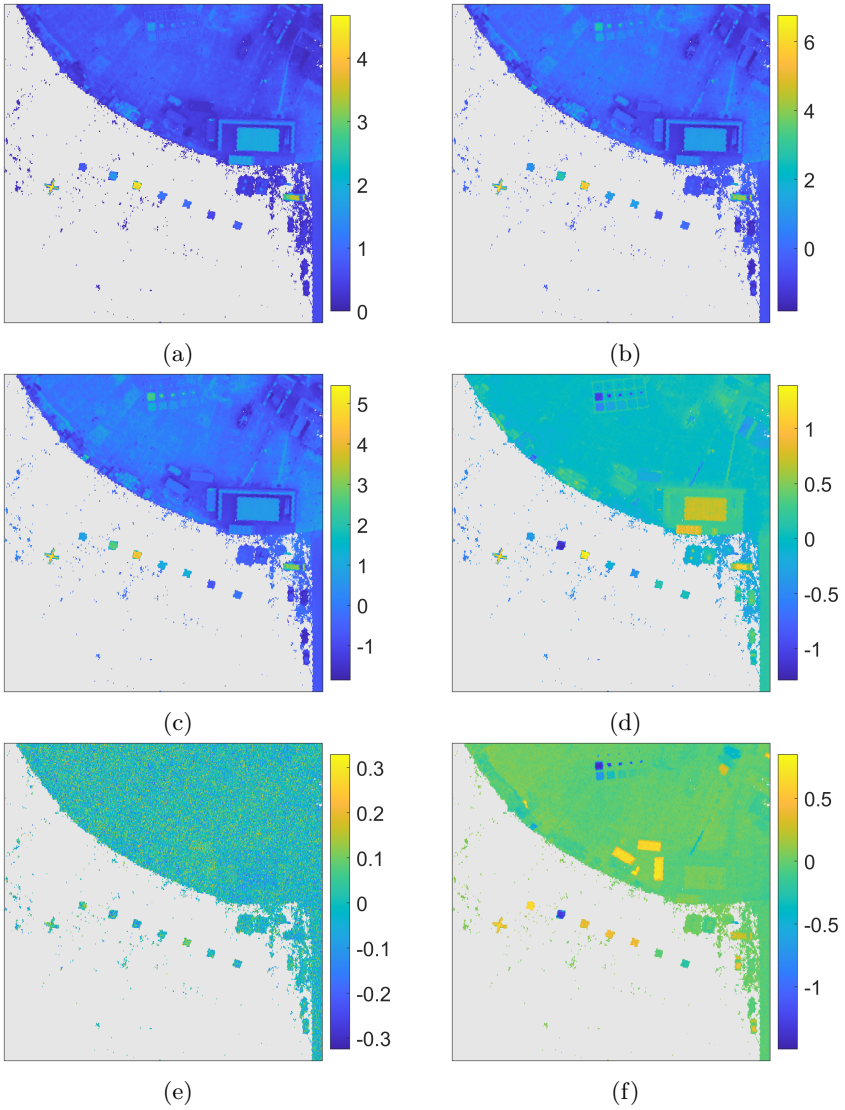


Figure 5.12: Principal Component Images of the UDRC Dataset. **a)** PC1 without background removal. **b)** PC1 with background removal. **c)** PC2 without background removal. **d)** PC2 with background removal. **e)** PC3 without background removal. **f)** PC3 with background removal.

5.4 Selection of the Optimal Target Detector

Each of the detection algorithms used were individually tested for their suitability when combined with the spatial and spectral DR schemes considered in the JSSDR-TD pipeline. In order to validate which algorithm performed optimally, the proposed JSSDR-TD method was applied to a subset of the UDRC Selene data. First, an ROC analysis was performed with examples of ROC curves for each combination of TD and DR algorithms are shown in Figure 5.13 for the full spatial scene and in Figure 5.14 when combined with NDVI_{re}.

As discussed in Section 3.6.7, ROC analysis, in isolation, is insufficient for comparing imbalanced binary classifiers [235]. However, it is interesting to note the disparity between the ROC curves produced using each of the TD algorithm outputs. In Figure 5.13 each of the algorithms used have near ideal ROC curves regardless of which spectral DR scheme is used when working on the full spatial scene. When spatial DR is employed, the AUC of each ROC curve, regardless of which combination of TD method and spectral DR algorithm is used, increases. In particular, both the ACE and CEM achieve near perfect ROC curves, as shown in Figure 5.14.

The AUC of the ROC curves increase for each spectral DR scheme when combined with NDVI_{re}-based spatial DR and the ACE algorithm, Figure 5.14a. By simplifying the background, and therefore improving the covariance estimate, the ACE algorithm achieved a better separation between the known target and the estimated background. Similarly, by suppressing the background, the FIR filter weight estimation that is necessary for the CEM algorithm is simplified. This is reflected in the increased AUC values of the ROC curve when using CEM with NDVI_{re}-based spatial DR, Figure 5.14b.

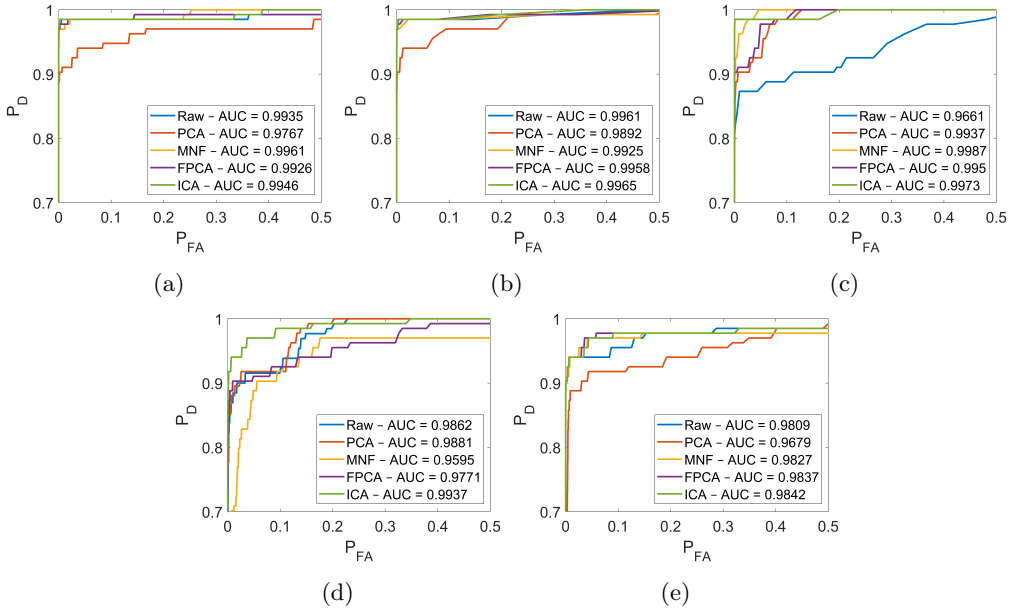


Figure 5.13: ROC Curves for each TD and spectral DR scheme pairing on the full scene. **a)** ACE. **b)** CEM. **c)** SAM. **d)** SID. **e)** RXD.

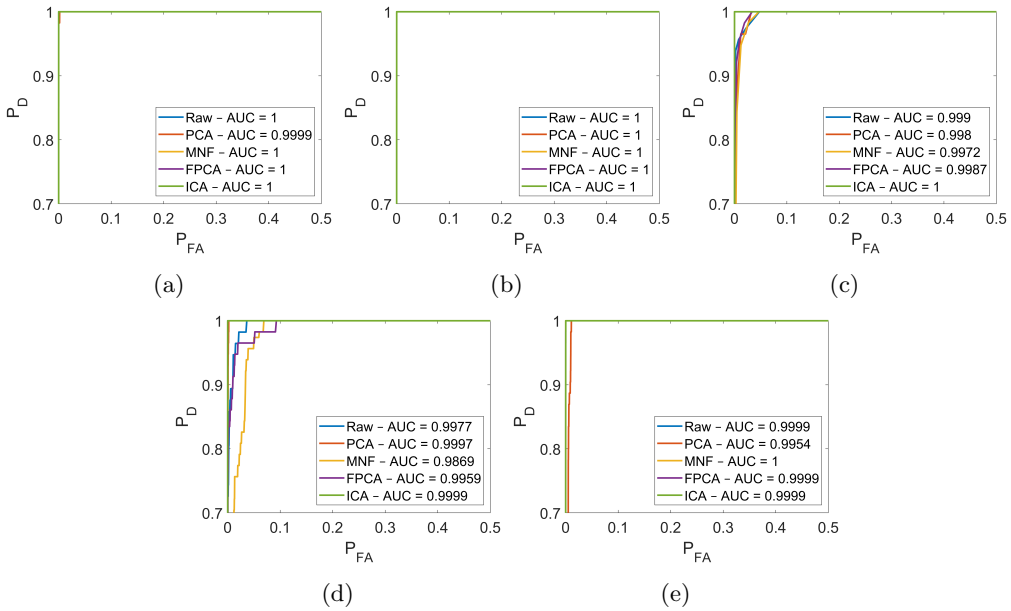


Figure 5.14: ROC Curves for each TD and spectral DR scheme pairing in combination with spatial DR. **a)** ACE. **b)** CEM. **c)** SAM. **d)** SID. **e)** RXD.

As well as ROC curves, PR curves were generated for each of the combinations of TD and DR algorithms with and without the $NDVI_{re}$ spatial DR. These are shown in Figure 5.15 when considering the full spatial scene and in Figure 5.16 when utilising $NDVI_{re}$-based spatial DR.

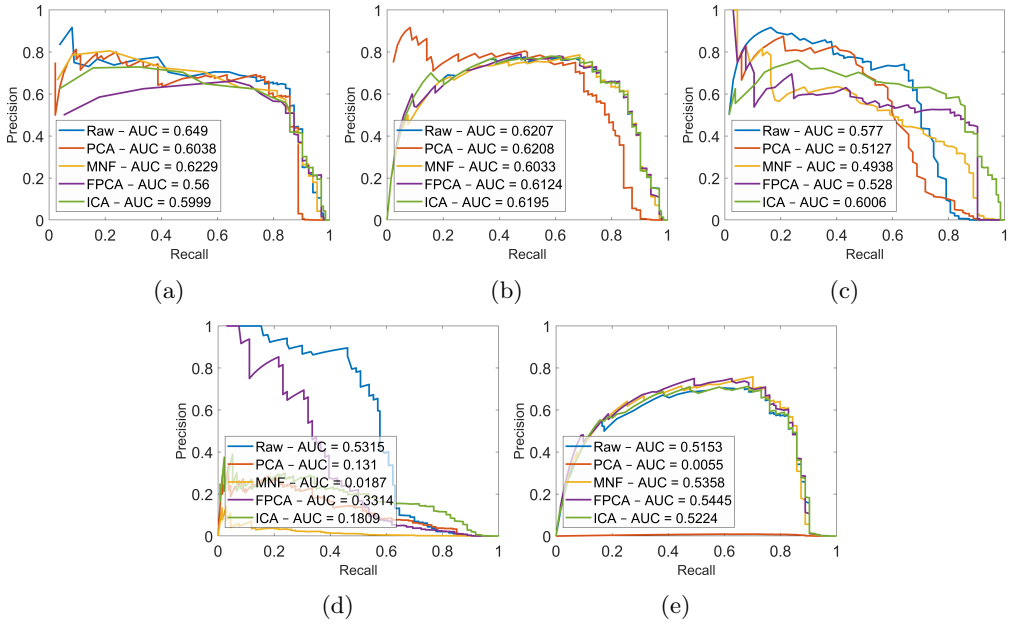


Figure 5.15: PR Curves for each TD and spectral DR scheme pairing on the full scene. **a)** ACE. **b)** CEM. **c)** SAM. **d)** SID. **e)** RXD.

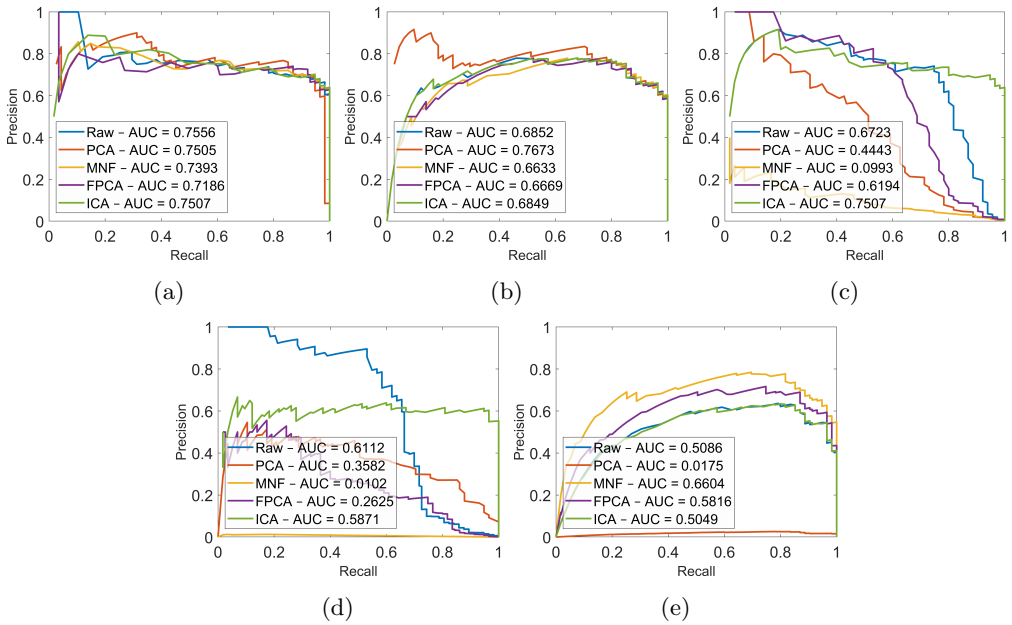


Figure 5.16: PR Curves for each TD and spectral DR scheme pairing in combination with spatial DR. **a)** ACE. **b)** CEM. **c)** SAM. **d)** SID. **e)** RXD.

Investigating the PR curves when considering the full scene, as shown in Figure 5.15, ACE, CEM and SAM all exhibit high AUC values for each of the spectral DR schemes used. SID, when used on the full dimensionality raw data can provide high AUC results and the RXD provides

average AUC using every spectral DR scheme, with the exception of PCA. As discussed, this can be attributed to the RXD being mathematically inverse to PCA and therefore unsuitable for use [152, 153]. When combined with spatial DR, Figure 5.16, the ACE, CEM and SAM algorithms continue to provide high AUC values. When using SID, the AUC of most methods, with the exception of FPCA and MNF, are increased. The AUC when using RXD remains largely the same following the application of spatial DR except when combined with MNF where there is an increase of approximately 10% AUC. The AUC of the PR curves are generally improved when using the $NDVI_{re}$-based spatial DR in conjunction with ACE, CEM and SAM.

Both the ROC and PR analysis were performed on a per-target basis. The results from Figures 5.13 to 5.16 are from the detection of a single target and are generally representative of the performance over every target present in the scene. Along with the ROC and PR curves, the other performance measures detailed in Section 3.6.7 were calculated for each of the targets in the scene. These measures were then averaged in order to obtain an overview of each TD algorithm's general performance, the results of which are presented in Figure 5.17.

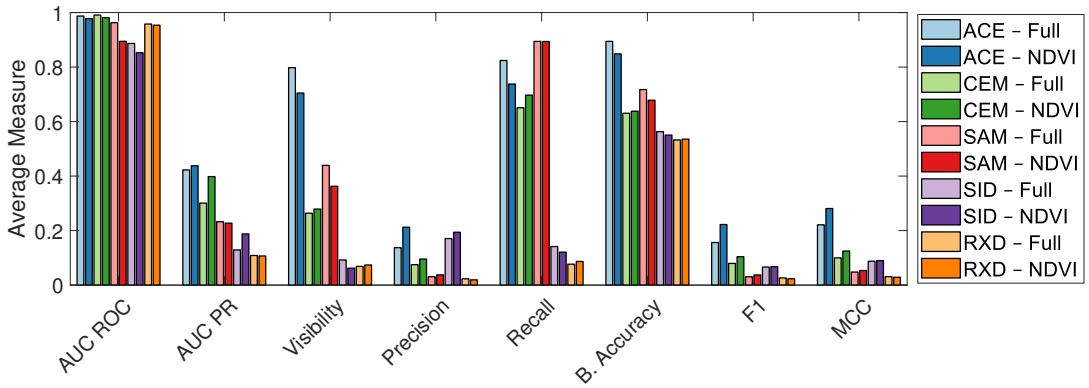


Figure 5.17: Average performance of each TD and DR algorithm combination when used on the full scene vs. when combined with spatial DR.

Similarly to the results drawn from Figures 5.13 and 5.14, each of the TD algorithms perform well when considering the AUC of the ROC curves. ACE and CEM give the highest AUC of the PR curves, with the performance of ACE improving when combined with spatial DR. Generally, using the spatial DR reduces the visibility of targets with the exception of CEM and the RXD where it slightly increases. ACE gives the highest visibility when considering both the full scene and when using spatial DR indicating that, of the algorithms investigated, it is the best at separating the background from the target. ACE and SID display the highest average precision, with both methods improving when using spatial DR. ACE also produces the highest balanced accuracy, F1 score, and MCC of each of the TD algorithms tested. For these

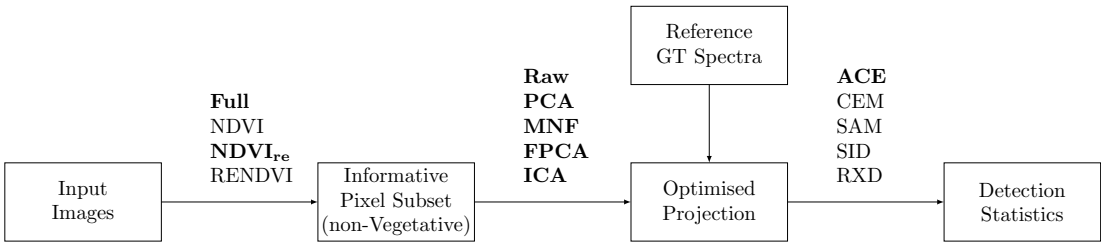


Figure 5.18: Algorithms used to test the proposed JSSDR-TD pipeline.

reasons, the remaining results in this chapter are generated using the ACE algorithm solely. In addition to reducing the sample size for increased efficiency, from Figure 5.17, for each of the detection algorithms, the performance after the application of spatial DR is generally as good or an improvement over using the full scene.

Figure 5.18 shows the pipeline from Figure 5.1 with the algorithms selected for use in each of the three stages, as discussed in Sections 5.2 to 5.4; spatial masking, spectral dimensionality reduction, and target detection, highlighted in bold. The results of performing TD using each combination of algorithms within the JSSDR-TD pipeline are presented in Section 5.5.

5.5 Results

5.5.1 Joint Spatio-Spectral Compression Performance

In this section, the compression rate and reduction in image dimensionality when using the combination of spatial and spectral DR within the JSSDR-TD are presented. $NDVI_{re}$ is used as a spatial mask, selecting pixels that are relevant and can be used in subsequent spectral DR and TD processes. By masking certain pixels they can be discarded from further processing, reducing the sample size. Then, by performing spectral DR, retaining K components from L spectral bands, where $K << L$, the sample size is reduced further. By combining the remaining spatial and spectral components, a compressed representation of the relevant data is retained for further processing. Table 5.2 details the size of each of the images used in this chapter, as well as their compressed spatial and spectral sizes along with the percentage of the original data retained after compression.

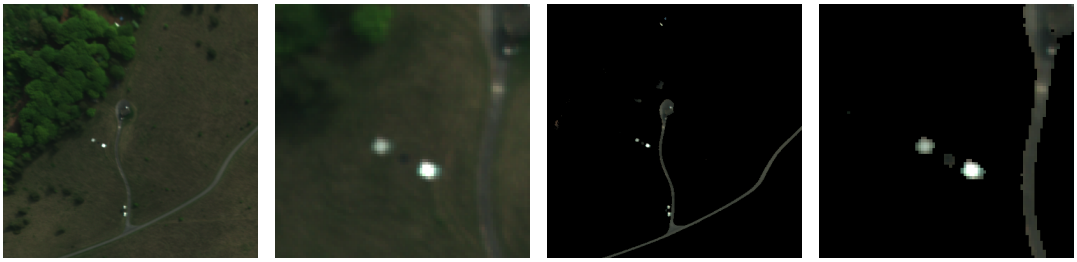
Table 5.2: Compression rates for combined spatial and spectral DR

Image	# Samples Full	# Samples $NDVI_{re}$ masked	L	K	Spatial Compressed Size (%)	Spectral Compressed Size (%)	Total Compressed Size (%)	Average Compressed Size (%)
OP7_1	160000	3504	100	20	2.19	20	0.44	
OP7_2	160000	2500	100	20	1.56	20	0.31	0.34
OP7_3	160000	2232	100	20	1.40	20	0.28	
IM_140804	3210191	649435	80	20	20.23	25	5.06	
IM_140806	3839976	578674	80	20	15.07	25	3.77	
IM_140807	3415052	689245	80	20	20.18	25	5.05	
IM_140808	3015944	543569	80	20	18.02	25	4.51	4.61
IM_140812	4360159	610172	80	20	13.99	25	3.50	
IM_140813	3301404	807262	80	20	24.45	25	6.11	
IM_140815	3640769	626776	80	20	17.22	25	4.30	

The OP7 dataset images are compressed to 1.72% of their original spatial extent on average as $NDVI_{re}$ selects a small proportion of the total pixels to process further. By retaining $K = 20$ components in the subsequent spectral DR stage, this is reduced further to an average of 0.34% of their original size. The images in the Selene trial have a much higher spatial resolution and a larger sample is retained after using the $NDVI_{re}$ spatial mask as a large proportion of the pixels represent non-target and non-vegetative materials, as shown in Figure 5.2. The pixels retained after $NDVI_{re}$ represent an average of 18.45% of the original image and applying spectral DR reduces this to 4.61% on average.

5.5.2 Results on the OP7 Dataset

The first of two datasets used in this chapter was provided by BAE Systems, Appendix A.3.1. It consists of three images of a forest scene, each capturing a common target area from a different flightpath. The target area contains three calibration panels, one grey, one black, and one white, which were used as the targets of interest. Figure 5.19a shows a false colour representation of one of the images with all three targets present in the scene. Figure 5.19c shows the same image masked using stage one of the JSSDR-TD pipeline from Figure 5.18, *i.e.,* NDVI_{re}-based spatial DR. Figures 5.19b and 5.19d are enlarged views of the target areas of Figures 5.19a and 5.19c respectively.



(a) (b) (c) (d)

Figure 5.19: Example of the OP7 Dataset. **a)** False colour image of the target area. **b)** Enlarged version of a).**c)** Retained pixels following the NDVI_{re}-based spatial masking. **d)** Enlarged version of c).

Of the two datasets considered in this chapter, OP7 is simpler as it contains fewer distinct materials and objects in the scene compared to the images from the UDRC Selene Trial dataset. The OP7 images also have a lower spatial resolution when compared to the Selene Trial data, with a Ground Sample Distance (GSD) of approximately 1 metre. As a result, approximately nine pixels per target contain pure spectra.

In order to assess how each TD algorithm's behaviour varied with the number of components retained using each DR scheme, the F1 score, MCC, balanced accuracy and visibility were calculated at various values for K, where $10 \leq K \leq L$ and $L = 100$ for the OP7 data. The average metrics for the three targets from Figures 5.19a and 5.19c are shown in Figure 5.20.

Both balanced accuracy and visibility are largely invariant to the number of components, K, retained, as shown in Figure 5.20. Although interestingly, at lower values of K the average target visibility using each spectral DR method is greater than that achieved using the raw data. Conversely, the F1 score and MCC both increase as the number of components increase to be equal to the original spectral dimensionality of the data, both with and without the application of spatial DR. This is to be expected, as in the case where $K = L$ the data is functionally

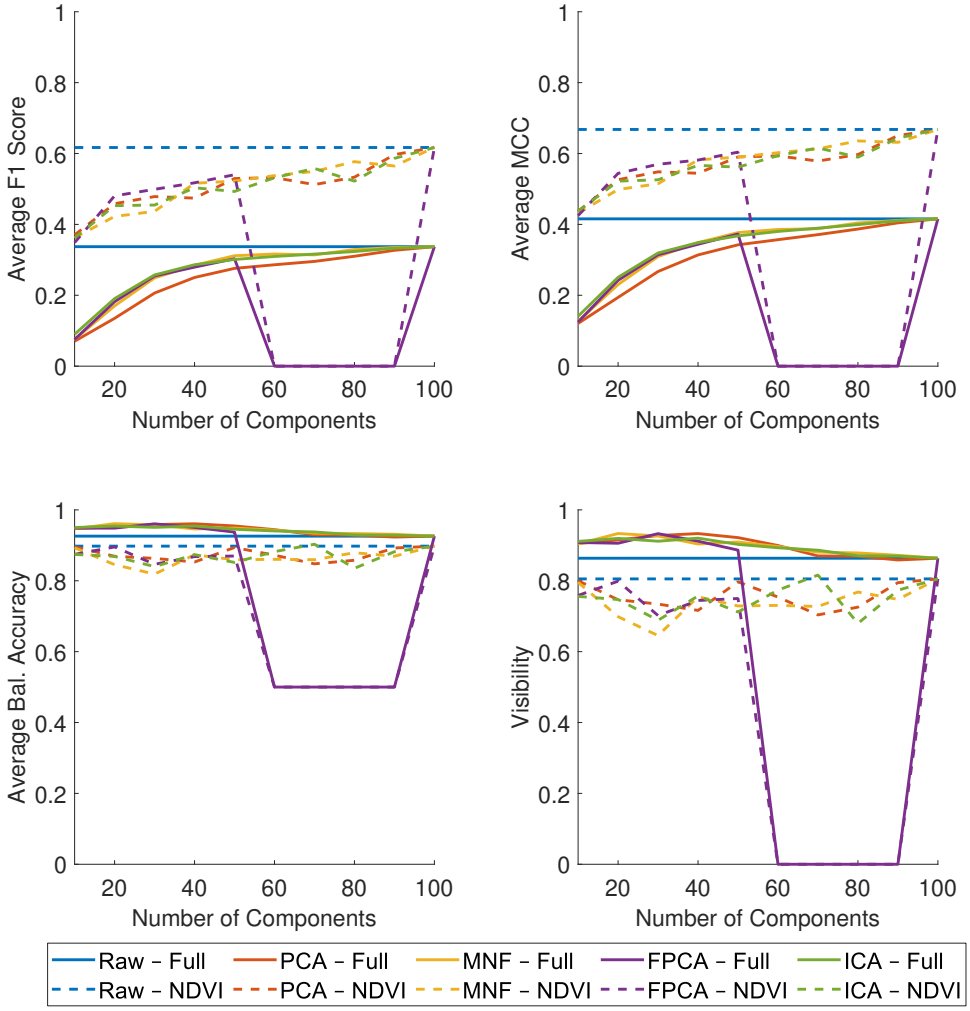


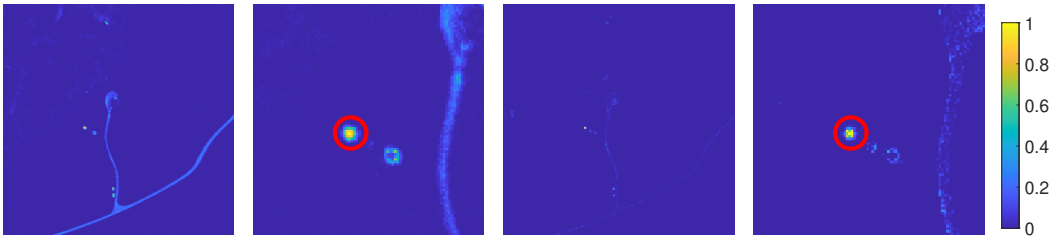
Figure 5.20: Performance measures using the ACE TD algorithm and each DR scheme with a varying number of retained components when detecting Target 1 (grey tile).

identical to the raw data and, although it has been remapped, no information has been lost. It can be noted that by using spatial DR prior to spectral DR, both the F1 score and MCC are increased beyond that achieved using the raw full dimensionality data without spatial DR. The FPCA folding parameters, H and W, were selected as factors of the number of components retained, K. This resulted in only certain values of K producing valid spectral DR results, as discussed further in Section 5.6.

When comparing TD performance on the full spatial dimensionality images with that of the NDVI_{re} masked images, both with and without the application of spectral DR, the F1 score and MCC both increase. However, when spatial DR is applied, the average F1 score and MCC are considerably higher. This is caused by the removal of the vegetative background which removes

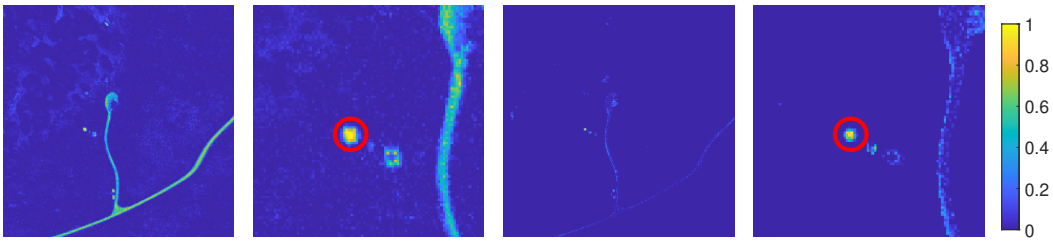
highly disparate observations and simplifies the problem of separating background from target. This increases the precision of the detection as seen in Tables 5.3 and 5.4. By reducing the complexity of the background, the targets, although more similar to the remaining background, can be separated in the subspace more easily. The MCC, in comparison to the F1 score, is slightly higher in both spatial DR cases as it takes into account the correct identification of the TN class. The balanced accuracy drops slightly wherever NDVI_{re} masking is applied. As the balanced accuracy is the average of the TPR and True Negative Rate (TNR), the decrease in the size of the TN class, without a corresponding proportional decrease in FPs, results in a lower balanced accuracy. Despite the increase in FPR when using NDVI_{re} masking, the absolute number of FP detections decreases. It can also be seen in Figure 5.20 that, by removing the easily separated vegetative background using NDVI_{re}, the visibility of the targets decrease, as only materials which, on average, are more spectrally similar to that of the target material remain.

Further comparisons were made by retaining $K = 20$ components from each of the spectral DR methods as this provided a good balance of performance and compression. From Figure 5.20, $K = 20$ components also gave clear improvements over the raw, full spatial and spectral dimensionality, scene when combined with spatial DR. The improvement in detection when using spatial DR can be seen in Figures 5.21 and 5.22, where the detection maps for target 1 (grey tile) are shown for both the full dimensionality image and following PCA-based spectral DR with and without NDVI_{re}-based spatial DR. Comparing Figures 5.21b and 5.21d there is less confusion in the detection map where NDVI_{re}-based spatial DR is applied. Similarly, using PCA with $K = 20$ components produces a similar effect, as shown in Figures 5.22b and 5.22d. The target is the brightest object in the scene in each case, indicating good separability between the target and the background with a high P_{d}.



(a) (b) (c) (d)

Figure 5.21: Detection statistics of the ACE algorithm detecting target 1 in the raw full dimensionality data where yellow indicates a target response of 1 and blue represents a target response of 0. **a)** Without spatial DR. **b)** Enlarged version of the target area in a). **c)** With NDVI_{re} applied. **d)** Enlarged version of the target area in c).



(a) (b) (c) (d)

Figure 5.22: Detection statistics of the ACE algorithm detecting target 1 in the the PCA projected data where $K = 20$. **a)** Without spatial DR. **b)** Enlarged version of the target area in a). **c)** With NDVI_{re} applied. **d)** Enlarged version of the target area in c).

In order to quantify this improvement, the ROC and PR curves for both the full and spatial dimensionality reduced images are shown in Figure 5.23 for each spectral DR method where $K = 20$. The F1 score, MCC and balanced accuracy when using each combination of spatial and spectral DR are plotted against the detection threshold in Figure 5.24.

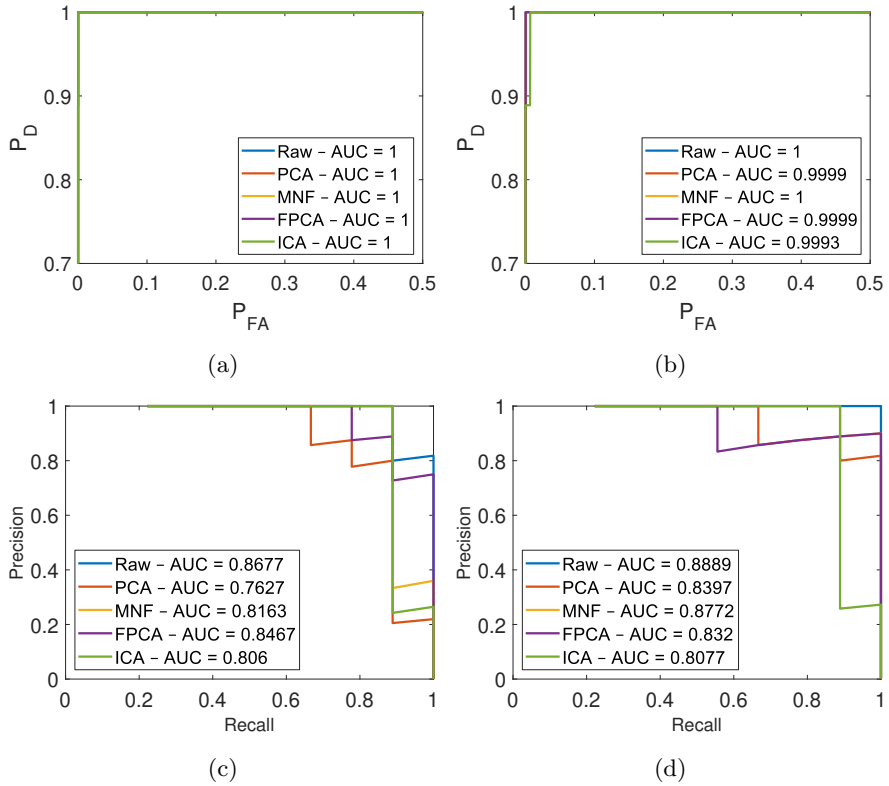


Figure 5.23: ROC and PR curves for detecting Target 1 (grey tile) in the OP7 dataset **a)** ROC curve using the full spatial dimensionality data. **b)** ROC curve using spatial DR pre-processing. **c)** PR curve using the full spatial dimensionality data. **d)** PR curve using spatial DR pre-processing.

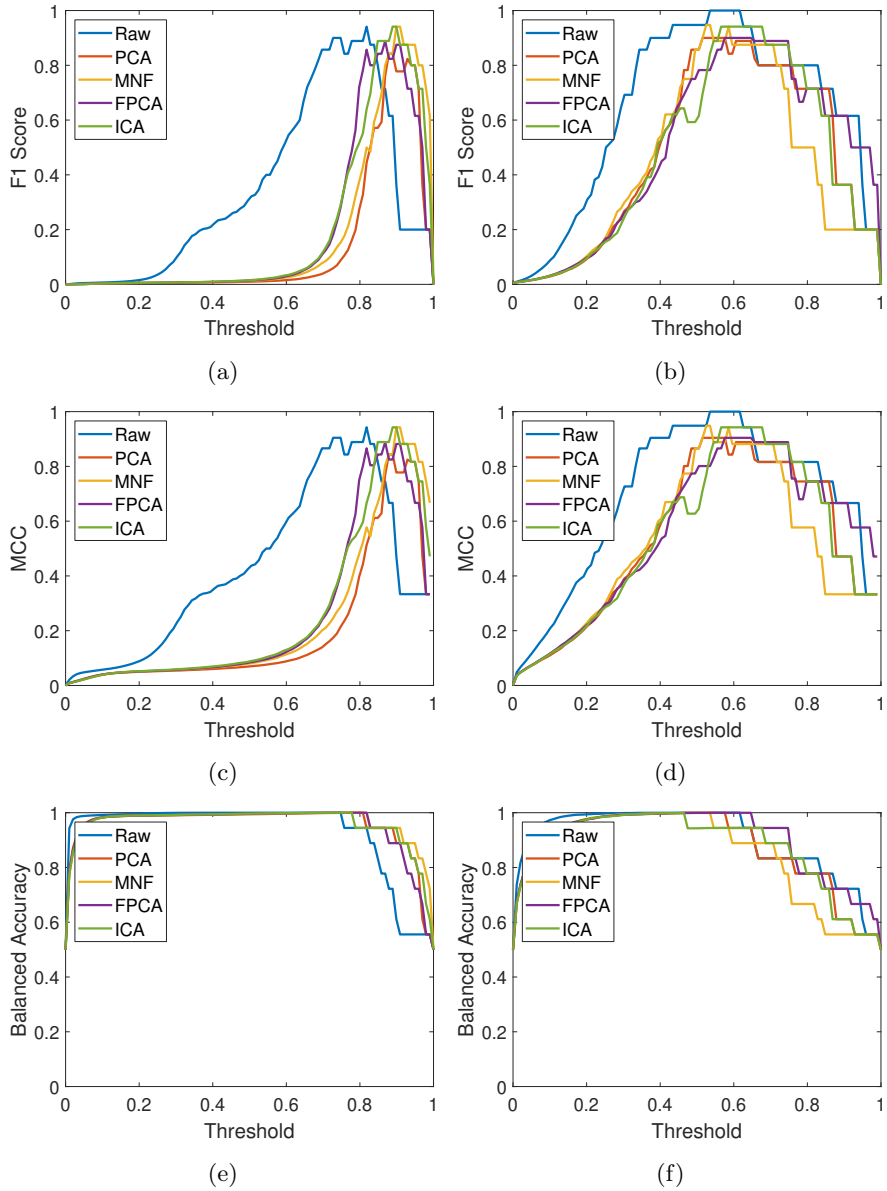


Figure 5.24: F1 score, MCC and balanced accuracy measures when detecting Target 1 (grey tile) at each detection threshold. **a)** F1 score when considering the full spatial scene. **b)** F1 score when considering the scene with spatial DR applied. **c)** MCC when considering the full spatial scene. **d)** MCC when considering the scene with spatial DR applied. **e)** Balanced accuracy when considering the full spatial scene. **f)** Balanced accuracy when considering the scene with spatial DR applied.

The ROC curves in Figures 5.23a and 5.23b are produced using the ACE detection statistics on the full scene and $NDVI_{re}$ masked scene respectively. Comparing the two sets of ROC curves is not informative, as the AUC is not significantly impacted by the application of $NDVI_{re}$-based spatial DR. The AUC of the PR curves (Figures 5.23c and 5.23d) when using

the raw uncompressed data, PCA or MNF dimensionality reduced data increases when spatial DR is applied. However, when applying FPCA or ICA, the AUC of the PR curve are not significantly altered.

Comparing the F1 scores and MCC from Figures 5.24a and 5.24b and Figures 5.24c and 5.24d respectively, it can be observed that for this target material (grey tile) a lower threshold value can be applied when using the JSSDR-TD pipeline to achieve the same scores as the image with spectral DR alone. Both the F1 score and MCC remain at a higher level over a larger range of threshold values when using spatial DR indicating a better separation between the target and the background. At higher thresholds both the F1 score and MCC decay as target pixels are excluded from the detection by the increasing threshold. This decay occurs at a slower rate when using spatial DR within the JSSDR-TD pipeline, allowing for more flexibility when selecting a detection threshold. The balanced accuracy measures behave similarly in the two scenarios as seen in Figures 5.24e and 5.24f. When considering the balanced accuracy measure in isolation, the full image scene allows for a wider range of thresholds that maximise this measure.

The results from Figures 5.20 to 5.24 are all calculated from a single target in order to display an example of the performance achieved. The average results for each target are shown in Table 5.3 when considering the full scene and in Table 5.4 when spatial DR has been applied.

Table 5.3: Average performance measures for each target in the OP7 dataset without spatial DR applied using the ACE algorithm.

K = 20	DR	ACE-Full							
		AUC ROC	AUC PR	Visibility	Precision	Recall	Bacc	F1	MCC
Grey Tile	Raw	**1.00**	**0.84**	0.88	**0.37**	0.88	0.93	**0.35**	**0.43**
	PCA	**1.00**	0.77	**0.94**	0.12	**0.95**	**0.96**	0.12	0.18
	MNF	**1.00**	0.79	0.93	0.15	**0.95**	**0.96**	0.15	0.21
	FPCA	**1.00**	0.80	0.92	0.17	0.94	**0.96**	0.17	0.23
	ICA	**1.00**	0.78	0.93	0.16	**0.95**	**0.96**	0.16	0.23
Black Tile	Raw	**1.00**	0.06	0.60	**0.09**	0.62	0.80	0.06	0.11
	PCA	**1.00**	0.10	0.68	0.05	0.72	0.84	0.05	0.09
	MNF	**1.00**	**0.15**	0.70	**0.09**	0.75	**0.85**	**0.07**	**0.12**
	FPCA	**1.00**	0.13	**0.71**	0.05	**0.76**	**0.85**	0.05	0.09
	ICA	**1.00**	0.11	0.67	0.04	0.72	0.83	0.05	0.09
White Tile	Raw	**1.00**	**0.74**	0.79	**0.57**	0.79	0.89	**0.53**	**0.59**
	PCA	**1.00**	0.67	**0.93**	0.22	**0.94**	**0.96**	0.28	0.37
	MNF	**1.00**	0.68	0.85	0.39	0.86	0.92	0.41	0.47
	FPCA	**1.00**	0.60	0.83	0.33	0.84	0.91	0.35	0.41
	ICA	**1.00**	0.67	0.79	0.47	0.80	0.89	0.44	0.49
All Spectra	Raw	**1.00**	**0.55**	0.76	**0.34**	0.77	0.87	**0.32**	**0.37**
	PCA	**1.00**	0.52	**0.85**	0.13	**0.87**	**0.92**	0.15	0.22
	MNF	**1.00**	0.54	0.83	0.21	0.85	0.91	0.21	0.27
	FPCA	**1.00**	0.51	0.82	0.18	0.85	0.91	0.19	0.25
	ICA	**1.00**	0.52	0.80	0.22	0.82	0.89	0.22	0.27

Table 5.4: Average performance measures for each target in the OP7 dataset with $NDVI_{re}$-based spatial DR applied using the ACE algorithm.

K = 20	DR	ACE-$NDVI_{re}$							
		AUC ROC	AUC PR	Visibility	Precision	Recall	Bacc	F1	MCC
Grey Tile	Raw	**1.00**	**0.86**	0.72	**0.73**	0.74	0.86	**0.59**	**0.65**
	PCA	**1.00**	0.83	0.75	0.56	0.79	0.87	0.48	0.54
	MNF	**1.00**	0.85	0.76	0.56	0.80	0.87	0.48	0.54
	FPCA	**1.00**	0.84	**0.81**	0.54	**0.84**	**0.90**	0.51	0.57
	ICA	**1.00**	0.75	0.73	0.52	0.77	0.86	0.43	0.50
Black Tile	Raw	**0.98**	**0.37**	**0.52**	**0.40**	**0.57**	**0.76**	**0.25**	**0.33**
	PCA	0.94	0.08	0.48	0.06	0.56	0.74	0.09	0.14
	MNF	0.96	0.09	0.46	0.07	0.54	0.73	0.09	0.14
	FPCA	0.94	0.09	0.47	0.08	0.54	0.73	0.10	0.15
	ICA	0.93	0.08	0.42	0.05	0.49	0.71	0.08	0.12
White Tile	Raw	**0.97**	**0.66**	0.57	**0.83**	0.58	0.78	0.58	**0.63**
	PCA	0.95	0.61	0.59	0.78	0.60	0.79	**0.59**	**0.63**
	MNF	0.95	0.62	**0.61**	0.74	**0.62**	**0.80**	0.58	0.62
	FPCA	0.92	0.59	0.59	0.64	0.61	0.79	0.51	0.55
	ICA	0.94	0.63	0.57	0.73	0.59	0.78	0.53	0.58
All Spectra	Raw	**0.98**	**0.63**	0.61	**0.65**	0.63	0.80	**0.47**	**0.53**
	PCA	0.96	0.50	0.61	0.47	0.65	0.80	0.39	0.44
	MNF	0.97	0.52	0.61	0.46	0.65	0.80	0.38	0.44
	FPCA	0.96	0.50	**0.62**	0.42	**0.66**	**0.81**	0.37	0.43
	ICA	0.96	0.49	0.57	0.43	0.62	0.78	0.35	0.40

The average results from Tables 5.3 and 5.4 are represented in a bar graph in Figure 5.25. In general, the AUC of the ROC and PR curves are similar regardless of the spectral and spatial DR used, with the exception of the full spectral dimensionality data where the AUC of the PR curve increases when using spatial DR. As discussed, and shown in Figure 5.20, the visibility drops significantly when using spatial DR as the highly dissimilar vegetative background is removed making the average background and target spectra more similar. The precision produced using the JSSDR-TD pipeline is increased when compared to those achieved using spectral DR alone. This is due to the reduction in FP detections as a result of the NDVI_{re} step masking dark areas of vegetation (which can be misclassified as the similarly dark black calibration tile *i.e.,* target 2) as well the improvement in separability following the spectral DR stage. The recall decreases slightly, however this may not be significant in TD applications as one pixel on target is sufficient for the identification and classification of an object of interest. Both the F1 score and MCC are improved when using spatial DR. In nearly all of the measures tested, the full spectral dimensionality image with and without spatial DR performed the best, the application of spatial DR was an improvement on average. Each of the spectral DR methods employed retained only 20 components of the original 100, reducing the computational complexity and cost of performing the TD while maintaining similar performance.

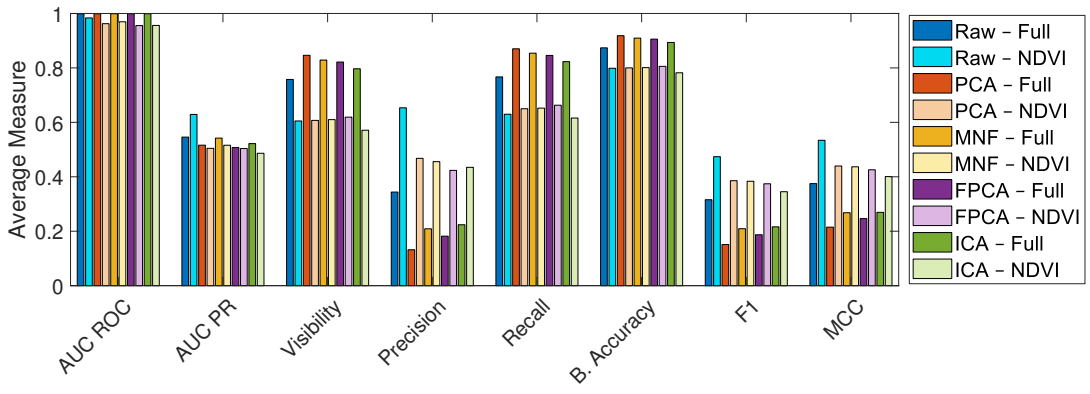


Figure 5.25: Comparison between the performance measures when combining spatial and spectral DR on the OP7 dataset

5.5.3 Results on the UDRC Selene Dataset

The second of the two datasets used in this chapter was provided by DSTL. It consists of seven images of a different forest scene with a large concrete area with metal containers, vehicles and other objects captured over the course of two weeks in August 2014. Each image covers a different view of this common target area containing between five and seven calibration panels of various colours and materials with a GSD of roughly 0.3 m. The targets present in each image are shown in Table 5.5. Figure 5.26a shows a false colour representation of the first of the images used (Image 1 in Table 5.5) where there are five of the possible seven target materials present in the scene. Figure 5.26c shows the image masked using NDVI_{re}-based spatial DR from stage one of the JSSDR-TD pipeline in Figure 5.1 with Figures 5.26b and 5.26d providing an enlarged view of the target area from Figures 5.26a and 5.26c respectively.

Table 5.5: Presence of target material in each image

Target	**IM140804**	**IM140806**	**IM140807**	**IM140808**	**IM140812**	**IM140813**	**IM140815**
Brown Carpet (T1)	×	×	×	×	×	×	×
Green Carpet (T2)	×	×	×	×	×	×	×
Green Ceramic (T3)	×	×	×	×	×	×	×
Green Perspex (T4)					×	×	×
Grey Ceramic (T5)	×	×	×	×	×	×	×
Orange Perspex (T6)				×	×	×	×
White Perspex (T7)	×	×	×	×	×	×	×

(a) (b) (c) (d)

Figure 5.26: Example of the UDRC Selene Dataset corresponding to Image 1 from Table 5.5. **a)** False colour image of the target area. **b)** Target region of a). **c)** Retained pixels following the NDVI_{re} spatial masking. **d)** Target region of c).

The same process of plotting the F1 score, MCC, balanced accuracy, and visibility of a target from the OP7 data against the number of components as in Figure 5.20 was applied to one of the target materials (target 3 - green ceramic), circled in Figures 5.26b and 5.26d. These graphs can be seen in Figure 5.27.

As in Figure 5.20 using the OP7 data, the average F1 score and MCC both increase with the number of retained components until $K = L$, in this case $L = 80$. ICA performs well

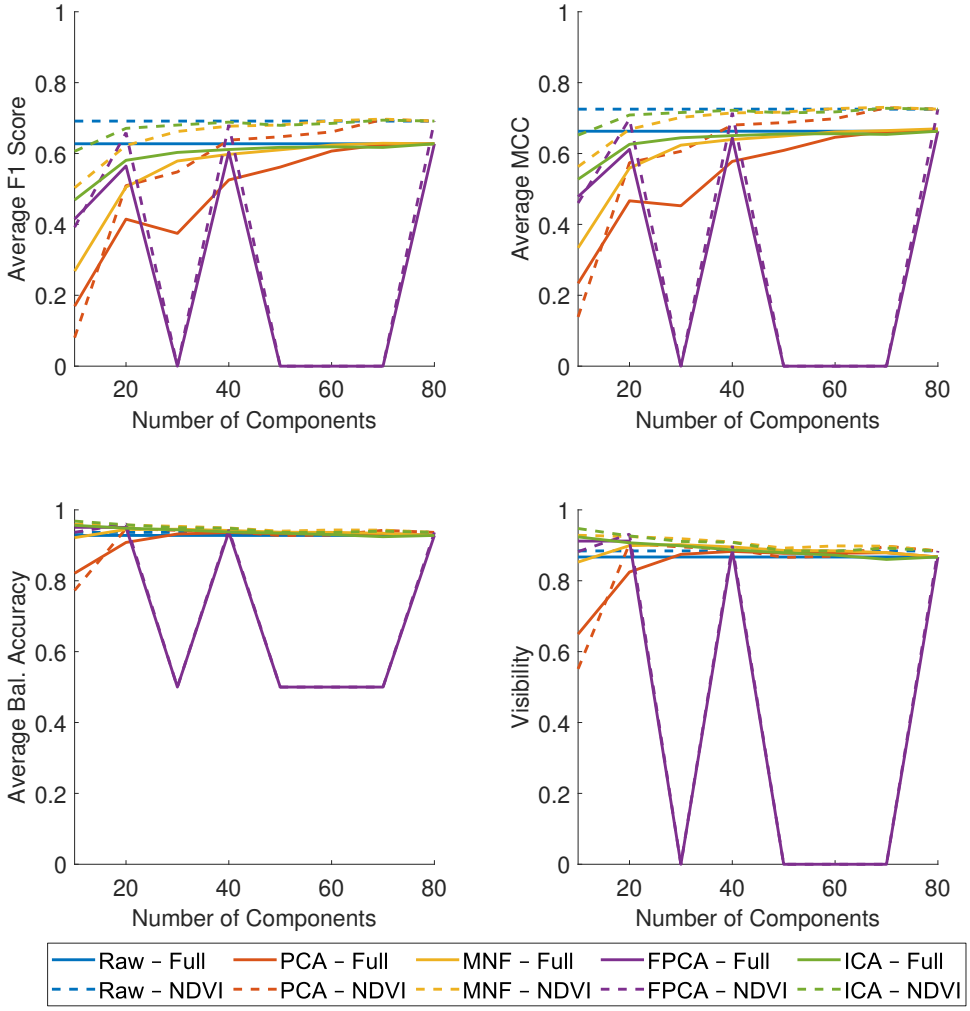
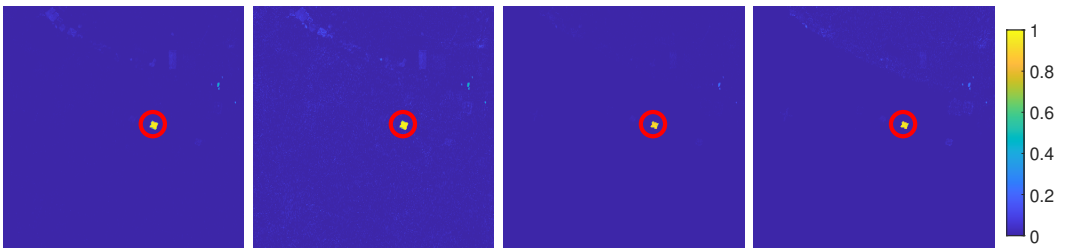


Figure 5.27: Performance measures using the ACE TD algorithm and each DR scheme with a varying number of retained components on the UDRC Selene dataset.

with $K = 20$ components, achieving a similar F1 score and MCC to that achieved using the raw data, with and without spatial DR. Both MNF and FPCA perform similarly to the raw data, although like the graphs displayed in Figure 5.20, only certain values of K are compatible with the folding parameters used in FPCA. PCA performs the worst in each of the metrics considered. Applying the spatial DR to the each of the spectral DR methods improves both their F1 score and MCC regardless of the number of components retained. Once again, the balanced accuracy and visibility are lowered when using spatial DR because of the reduced TN class and more similar average background signature respectively. Similar to the results gathered from the OP7 dataset, applying spectral DR improves the balanced accuracy and visibility of the full spatial scene at lower values of K. The remaining results were obtained



(a) (b) (c) (d)

Figure 5.28: Detection statistics of the ACE algorithm detecting target 1 in UDRC image using the following input data: **a)** raw, full dimensionality, image, **b)** Raw, full dimensionality, image with NDVI_{re}-based spatial DR, **c)** ICA projected image with $K = 20$, **d)** ICA projected image with $K = 20$ and NDVI_{re}-based spatial DR.

with $K = 20$, as it provided a good balance between detector performance and compression. The results given in Figure 5.27 also indicate that improved performance could be obtained at $K = 40$ at the expense of compression rate.

Similar to the results obtained on the OP7 dataset in Figure 5.21, removing the vegetation and simplifying the background class improves separation between the synthetic background and targets. While there is an overall decrease in target visibility, as the average spectra are more similar to the targets, there is less varied information to be represented, either in the full dimensionality image or in a dimensionality reduced subspace. This leads to less confusion in the detection image as shown in Figure 5.28.

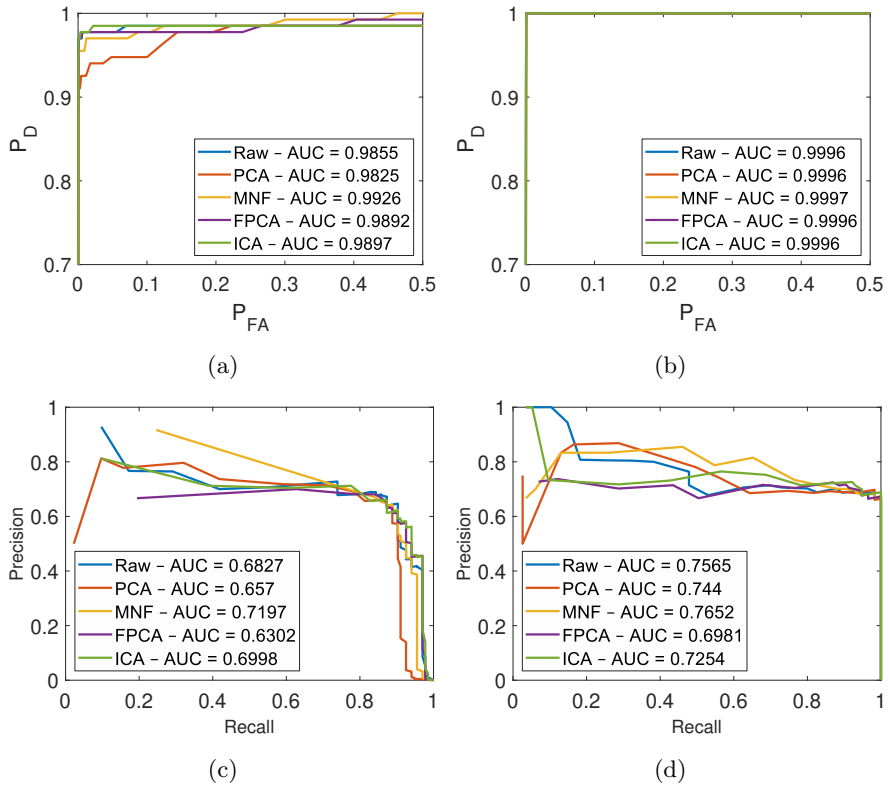


Figure 5.29: ROC and PR curves for detecting Target 3 (green ceramic) in the UDRC dataset **a)** ROC curve using the full spatial dimensionality data. **b)** ROC curve using spatial DR pre-processing. **c)** PR curve using the full spatial dimensionality data. **d)** PR curve using spatial DR pre-processing.

The ROC curves in Figures 5.29a and 5.29b are produced using the ACE detection statistics on the full scene and $NDVI_{re}$-based spatially masked scene respectively. The two sets of ROC curves are almost identical and do not provide definitive results, but indicate a small improvement when using the spatial DR. Comparing the PR curves in Figures 5.29c and 5.29d show that when each spectral DR scheme is used in conjunction with spatial DR the AUC is increased by 5-10%.

Similar to the results obtained from the OP7 dataset, when considering both the F1 score (Figures 5.30a and 5.30b) and MCC (Figures 5.30c and 5.30d), a lower threshold can be used to produce improved results with the application of $NDVI_{re}$-based spatial DR compared to those achieved using the full scene. However, unlike the results from Figure 5.20 the balanced accuracy of the ACE algorithm is improved at higher thresholds when spatial DR is used on the UDRC Selene Trial data, as shown in Figures 5.30e and 5.30f. There is less decay at higher thresholds, when compared with the results from Figure 5.20, indicating that this target is detected with more certainty as the response from the ACE algorithm is higher, Figure 5.28.

When comparing the performance of the ACE algorithm on the full spatial scene with that of the $NDVI_{re}$ masked scene, the F1 score and MCC are both higher over a slightly larger range of thresholds when spatial DR is used. With the exception of PCA, which requires a similar threshold to obtain similar performance, each method obtains higher a F1 score and MCC at lower thresholds. On average, the balanced accuracy is increased over the range of thresholds considered when spatial DR is applied.

The results from Figures 5.27 to 5.30 are all calculated from a single target material (green ceramic) in order to display an example of the performance achieved. The average results for each target in the Selene dataset are shown in Table 5.6 when considering the full scene and in Table 5.7 when using the JSSDR-TD pipeline. The average results from the detection of each target can be seen combined in Figure 5.31.

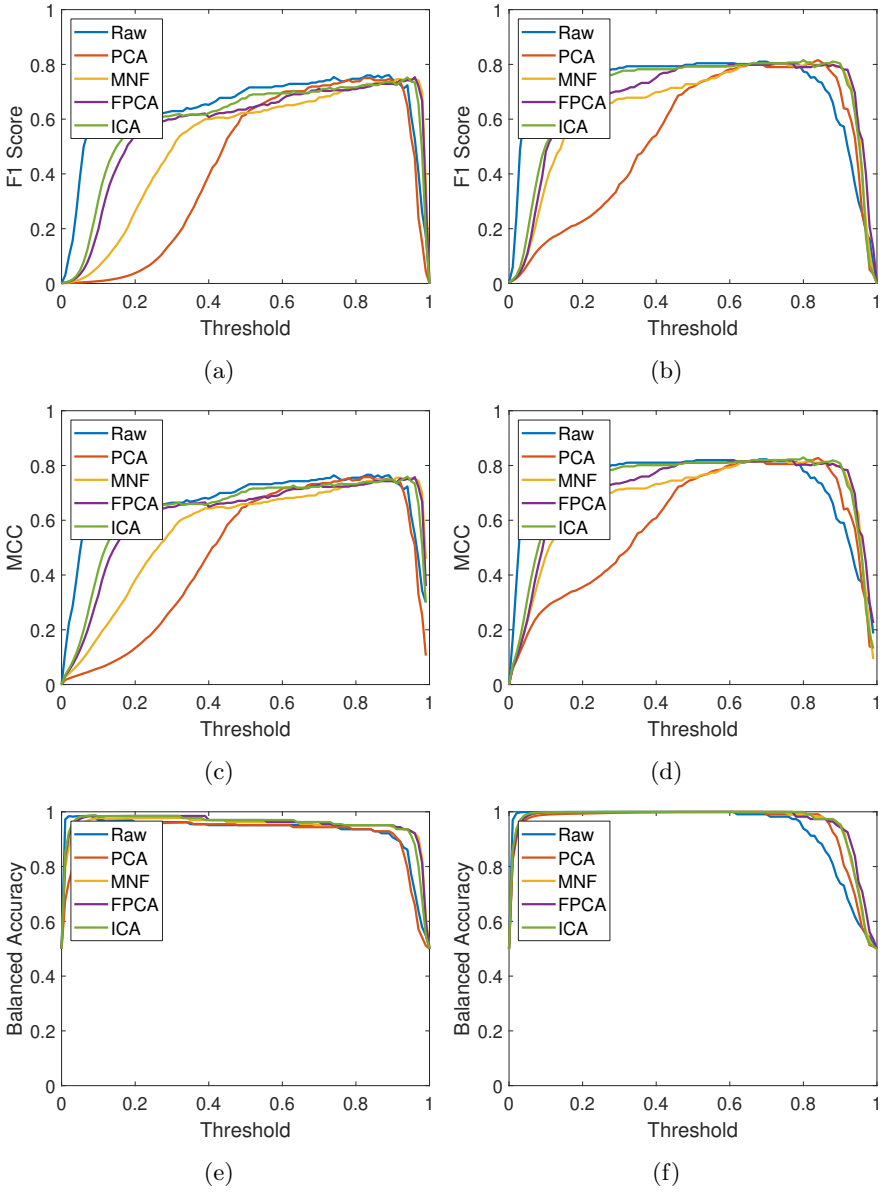


Figure 5.30: F1 score, MCC and balanced accuracy measures when detecting Target 3 (green ceramic) at each detection threshold. **a)** F1 score when considering the full spatial scene. **b)** F1 score when considering the scene with spatial DR applied. **c)** MCC when considering the full spatial scene. **d)** MCC when considering the scene with spatial DR applied. **e)** Balanced accuracy when considering the full spatial scene. **f)** Balanced accuracy when considering the scene with spatial DR applied.

Table 5.6: Average performance measures for each target in the Selene dataset without spatial DR applied using the ACE algorithm.

K = 20	DR	ACE-Full							
		AUC ROC	AUC PR	Visibility	Precision	Recall	Bacc	F1	MCC
Brown Carpet	Raw	**0.97**	0.54	0.64	**0.42**	0.65	0.81	**0.35**	**0.41**
	PCA	**0.97**	0.32	0.60	0.20	0.64	0.80	0.14	0.21
	MNF	**0.97**	0.56	0.71	0.27	0.75	0.85	0.25	0.31
	FPCA	**0.97**	0.59	**0.73**	0.31	0.76	**0.86**	0.28	0.35
	ICA	**0.97**	**0.61**	**0.73**	0.33	**0.78**	**0.86**	0.30	0.36
Green Carpet	Raw	**0.98**	**0.61**	0.79	**0.45**	0.79	0.89	**0.49**	**0.54**
	PCA	0.94	0.21	0.48	0.16	0.54	0.74	0.10	0.15
	MNF	0.96	0.54	0.70	0.32	0.74	0.85	0.30	0.35
	FPCA	**0.98**	0.59	0.83	0.34	0.86	**0.91**	0.39	0.45
	ICA	**0.98**	0.59	**0.84**	0.31	**0.88**	**0.91**	0.36	0.42
Green Ceramic	Raw	**0.99**	**0.69**	0.88	**0.55**	0.88	0.94	**0.64**	**0.67**
	PCA	**0.99**	0.67	0.88	0.36	0.90	0.94	0.44	0.50
	MNF	**0.99**	0.66	0.91	0.42	**0.93**	**0.95**	0.53	0.58
	FPCA	**0.99**	0.64	**0.92**	0.46	**0.93**	**0.95**	0.57	0.61
	ICA	**0.99**	0.64	0.91	0.46	0.92	**0.95**	0.56	0.60
Green Perspex	Raw	**1.00**	**0.65**	0.88	**0.53**	0.88	0.93	**0.62**	**0.66**
	PCA	**1.00**	0.61	**0.93**	0.36	0.93	**0.96**	0.48	0.54
	MNF	**1.00**	0.62	**0.93**	0.43	**0.94**	**0.96**	0.56	0.61
	FPCA	**1.00**	0.62	0.92	0.46	0.93	**0.96**	0.58	0.62
	ICA	**1.00**	**0.65**	0.92	0.46	0.93	**0.96**	0.57	0.62
Grey Ceramic	Raw	0.98	0.63	0.70	**0.46**	0.71	0.84	**0.40**	**0.46**
	PCA	0.98	0.58	0.71	0.32	0.74	0.85	0.26	0.33
	MNF	**0.99**	**0.66**	0.78	0.36	0.82	0.89	0.36	0.42
	FPCA	**0.99**	0.63	**0.79**	0.35	**0.83**	**0.89**	0.37	0.43
	ICA	**0.99**	0.61	0.76	0.33	0.80	0.88	0.33	0.39
Orange Perspex	Raw	**0.99**	**0.67**	0.85	**0.59**	0.85	0.92	**0.66**	**0.69**
	PCA	**0.99**	0.64	**0.91**	0.41	**0.92**	**0.95**	0.52	0.57
	MNF	**0.99**	0.65	**0.91**	0.49	0.92	**0.95**	0.59	0.64
	FPCA	**0.99**	0.62	**0.91**	0.52	0.91	**0.95**	0.62	0.66
	ICA	**0.99**	0.62	**0.91**	0.51	**0.92**	**0.95**	0.62	0.66
White Perspex	Raw	0.95	0.19	0.32	**0.17**	0.34	0.66	0.12	0.16
	PCA	**0.99**	**0.29**	**0.72**	0.13	**0.74**	**0.85**	**0.15**	**0.22**
	MNF	0.98	0.25	0.57	0.15	0.60	0.78	0.13	0.19
	FPCA	0.96	0.17	0.44	0.15	0.48	0.72	0.10	0.15
	ICA	0.90	0.12	0.32	0.10	0.36	0.66	0.07	0.11
All Spectra	Raw	**0.98**	**0.56**	0.70	**0.43**	0.71	0.85	**0.44**	**0.49**
	PCA	**0.98**	0.45	0.72	0.26	0.75	0.85	0.26	0.32
	MNF	**0.98**	**0.56**	**0.77**	0.33	**0.80**	**0.88**	0.35	0.41
	FPCA	**0.98**	0.54	**0.77**	0.35	**0.80**	**0.88**	0.38	0.44
	ICA	0.97	0.54	0.75	0.33	0.78	0.87	0.37	0.42

Table 5.7: Average performance measures for each target in the Selene dataset with $NDVI_{re}$-based spatial DR applied using the ACE algorithm.

K = 20	DR	ACE-$NDVI_{re}$							
		AUC ROC	AUC PR	Visibility	Precision	Recall	Bacc	F1	MCC
Brown Carpet	Raw	**0.99**	0.38	0.47	**0.36**	0.50	0.73	0.26	0.31
	PCA	0.92	0.08	0.30	0.07	0.37	0.65	0.06	0.10
	MNF	0.97	0.25	0.43	0.23	0.49	0.71	0.16	0.21
	FPCA	0.98	0.41	0.54	0.32	0.59	0.76	0.24	0.30
	ICA	**0.99**	**0.55**	**0.66**	0.35	**0.72**	**0.83**	**0.31**	**0.38**
Green Carpet	Raw	**1.00**	**0.63**	0.76	**0.50**	0.77	0.87	**0.51**	**0.56**
	PCA	0.90	0.08	0.25	0.07	0.31	0.63	0.06	0.09
	MNF	0.99	0.47	0.61	0.39	0.65	0.80	0.30	0.36
	FPCA	**1.00**	0.61	**0.79**	0.39	0.83	0.89	0.43	0.49
	ICA	**1.00**	**0.63**	0.85	0.39	**0.89**	**0.92**	0.45	0.51
Green Ceramic	Raw	**1.00**	**0.71**	0.84	**0.62**	0.84	0.91	**0.65**	**0.69**
	PCA	**1.00**	0.69	0.85	0.49	0.87	0.92	0.55	0.60
	MNF	**1.00**	0.66	0.90	0.53	0.91	0.94	0.61	0.66
	FPCA	**1.00**	0.65	0.90	0.55	0.91	0.94	0.63	0.67
	ICA	**1.00**	0.66	**0.91**	0.53	**0.93**	**0.95**	0.62	0.66
Green Perspex	Raw	**1.00**	**0.69**	0.83	**0.61**	0.83	0.91	0.63	0.67
	PCA	**1.00**	0.63	0.90	0.52	0.91	0.94	0.61	0.65
	MNF	**1.00**	0.64	0.90	0.55	0.91	**0.95**	**0.64**	**0.68**
	FPCA	**1.00**	0.63	0.90	0.56	0.90	0.94	**0.64**	**0.68**
	ICA	**1.00**	0.66	**0.91**	0.53	**0.93**	**0.95**	0.63	0.67
Grey Ceramic	Raw	**1.00**	**0.64**	0.62	**0.53**	0.64	0.81	**0.41**	**0.47**
	PCA	**1.00**	0.56	0.63	0.37	0.69	0.81	0.30	0.37
	MNF	**1.00**	0.62	0.69	0.39	0.74	0.84	0.34	0.41
	FPCA	**1.00**	**0.64**	**0.71**	0.40	**0.77**	**0.85**	0.37	0.43
	ICA	0.98	0.54	0.67	0.29	0.74	0.83	0.29	0.35
Orange Perspex	Raw	**1.00**	**0.69**	0.72	**0.63**	0.72	0.85	0.60	0.63
	PCA	**1.00**	0.65	0.84	0.54	0.85	0.91	0.60	0.64
	MNF	**1.00**	0.66	0.85	0.57	0.86	0.92	0.64	0.67
	FPCA	**1.00**	0.65	0.82	0.57	0.82	0.90	0.63	0.66
	ICA	**1.00**	0.66	**0.87**	0.57	**0.88**	**0.93**	**0.65**	**0.68**
White Perspex	Raw	0.86	0.07	0.17	0.09	0.22	0.59	0.06	0.08
	PCA	**0.99**	**0.45**	**0.54**	**0.29**	**0.60**	**0.77**	**0.22**	**0.28**
	MNF	0.93	0.20	0.35	0.15	0.42	0.68	0.11	0.15
	FPCA	0.84	0.11	0.27	0.10	0.34	0.63	0.07	0.10
	ICA	0.80	0.06	0.20	0.08	0.28	0.60	0.05	0.08
All Spectra	Raw	0.97	**0.53**	0.60	**0.45**	0.62	0.80	**0.42**	**0.46**
	PCA	0.97	0.42	0.57	0.30	0.62	0.78	0.29	0.34
	MNF	**0.98**	0.48	0.64	0.37	0.68	0.82	0.36	0.41
	FPCA	0.97	0.51	0.68	0.39	0.72	0.83	0.39	0.44
	ICA	0.96	0.52	**0.70**	0.37	**0.74**	**0.85**	0.39	0.44

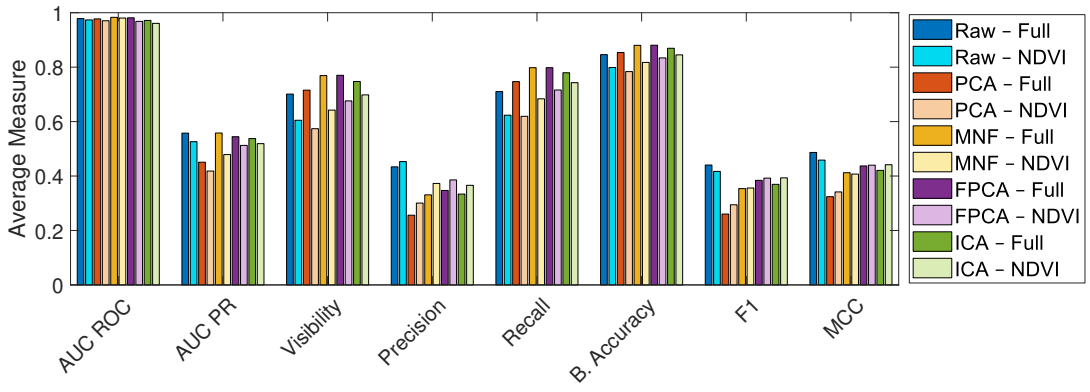


Figure 5.31: Comparison between the performance measures when combining spatial and spectral DR on the Selene dataset.

In general, from Figure 5.31 and Tables 5.6 and 5.7, the AUC of both the ROC and PR curves is similar regardless of the spectral and spatial DR used. As shown in Figure 5.27, there is a decrease in visibility of the target when using spatial DR as the highly dissimilar vegetative background is removed. This improves as the number of retained components, K, is increased at the expense of compression. The precision when using the JSSDR-TD pipeline increases in comparison to using only spectral DR on the full spatial scene, with the recall dropping slightly. Both the F1 score and MCC are improved when using spatial DR in all methods tested.

Similar to the results from the OP7 images, the full dimensionality images with and without spatial DR have the best performance. However, of the spectral DR methods used, both FPCA and ICA perform well despite retaining the equivalent of only 25% of the total spectral components. When combined with spatial DR both ICA and FPCA produce representations of the original data which maintain the same level of TD performance as the full dimensionality image with no spatial DR applied.

The performance measures when applying the JSSDR-TD to the Selene dataset (Figure 5.25 and Tables 5.3 and 5.4) offer an improvement in some metrics compared to using spectral DR alone. However, these improvements are not as significant as those achieved from the processing of the OP7 dataset. This can be attributed to the increased complexity of the Selene data when compared to the OP7 data. The performance can be improved further by increasing the number of components retained in stage 2 of the pipeline from Figure 5.1, as shown in Figure 5.27, albeit at the expense of compression and therefore at an increased computational cost.

5.6 Summary

In this chapter, the novel JSSDR-TD pipeline has been presented and used to improve the efficiency of aerial hyperspectral TD. The proposed $NDVI_{re}$-based spatial DR is relatively simple, requiring information from only two wavelengths and can be readily applied to TD and other similar applications. By using $NDVI_{re}$ it is possible to separate varied spectral targets composed of metals, plastics and other synthetic materials against a vegetative background. NDVI variants allow for the discrimination between vegetative and non-vegetative pixels due to known material characteristics in the NIR and red-edge portions of the spectrum. Other VIs, while not considered here, as exploiting the red-edge portion was determined to be a key component of this method, may provide alternative insights and allow for the more optimal detection of additional materials in alternative environments. Additionally, as the datasets tested were scenes of a predominantly vegetative environment, VIs were considered to be most useful as a method of performing spatial DR. Other indices may prove useful for performing similar processes in different environments such as the water and snow indices for removal of appropriate background information, or using the BUI for isolating and identifying man-made objects.

The complexity and performance of the spectral DR methods utilised varies. PCA is the simplest method used but also requires the most spectral components to be retained in order to be competitive. Applying the spatial DR and simplifying the background prior to performing spectral DR improved the performance of all methods but most notably when using PCA, which was competitive in both datasets with the addition of spatial DR. MNF can be conceptualised as two PCAs, one for noise reduction and the second to transform the noise whitened data into the reduced subspace. This extra noise removal step offers a distinct advantage when compared to PCA and allows it to perform similarly to FPCA and ICA.

FPCA performed favourably in both datasets and is efficient given the simplification when calculating the partial covariance matrix. However, when using FPCA the choices of the number of components, K, and the height, H, and width, W, of the folded matrix are far more sensitive than the other methods and are subject to two rules:

1. K must be a factor of the total number of wavelengths, L
2. When selecting the folding parameters H and W, $L > (H-1)W$

In any case where the first rule is true, the expression in the second rule will automatically be valid. H was selected to be half the value of K in order to adapt with the changing

number of components. However, due to the folding process, padding the folded array with zeros in order to fulfil the expression $HW = L$, if these zeros formed an entire row of the covariance matrix they will form a zero component in both the projected image and target. When these interact in each of the TD algorithms, usually by inner product, it forms a singular matrix. As inverse matrices are prevalent in the implementations of the TD algorithms used, singular matrices completely suppresses the detection. This phenomenon caused the undulating behaviours present in Figures 5.20 and 5.27 and informed the choice of the number of DR components retained when comparing each TD algorithm.

ICA is the most complicated and computationally expensive method to compute, but performed well on both datasets. Only using the full dimensionality data, with and without spatial DR was an improvement over the ICA-based methods. In general, the spectral DR methods, while increasing the balanced accuracy and visibility when smaller numbers of components are retained, decrease the F1 score and MCC when compared with the raw full dimensionality data. Both FPCA and ICA offer consistent and improved detection when combined with ACE and NDVI_{re}-based spatial DR. In general, the most impressive results are obtained using the ACE TD algorithm which corroborates the conclusions of other similar works investigating this topic [120, 121, 124].

The methods detailed here offered improvement to the TD performance on both datasets considered. However, greater improvements were obtained on the simpler OP7 dataset. Increasing the number of spectral DR components retained to account for the increased variability in the UDRC Selene dataset would improve the performance. This is at the expense of the compression rates that can be achieved at lower values of K. On average, applying NDVI_{re}-based spatial DR increases precision and slightly decreases the recall of the TD algorithms used. The visibilities of the targets decrease as background pixels which are dissimilar to the targets are not considered. The average background signature, after applying the NDVI_{re}-based spatial DR, becomes more similar to the target signatures. However, applying spectral DR and mapping the data into a more informative subspace can alleviate this issue.

By combining both spatial and spectral DR, the computational complexity and memory requirements can be reduced whilst maintaining, or in some cases improving upon, detection performance as shown in Figures 5.25 and 5.31. This is achieved using data compressed to $< 5\%$ of the total size the original data in the JSSDR-TD pipeline. Using spatial DR had little effect on the AUC of the ROC or PR curves, the main improvements came from the increased F1 score, MCC and precision. On average, there is a slight reduction in recall and balanced accuracy, however, this is likely due to the removal of the most obvious non-target pixels. While

a subset of results are shown in this chapter, additional results on alternative targets and images from both the OP7 and UDRC datasets can be found in Appendices B.5 and B.6.

Chapter 6

Sparsity and Compressive Sensing Techniques for Object Detection in Hyperspectral Imagery

6.1 Introduction

As discussed in Chapter 5, the detection and subsequent classification of objects in aerial imagery is integral to multiple applications. HSI offers a vast quantity of spectral information about a scene and, as sensor technology improves, the amount of spatial data available is also increasing. With this increasing clarity comes an increase in the computational overhead required to both obtain and analyse this data. Traditional spectral imaging techniques, discussed in Section 3.4.1, capture adjacent pixels in an image at the expense of spatial, spectral, or temporal resolution, or indeed as is more common, a trade-off between all three. Whilst DR-based techniques were investigated in Chapter 5, in this chapter Compressive Spectral Imaging and the exploitation of the sparsity exhibited by hyperspectral images will be investigated.

There are various applications in remote sensing which already benefit from the use of CS techniques, including Endmember estimation [282], object detection [283] and image fusion [284] to name a few. With the implementation of CS-based image capture in hardware using CASSI-based systems [184, 285] enabling CS in multiple imaging applications.

Coded Aperture Snapshot Spectral Imaging [184, 185] is an innovative CSI approach that captures a 3D spatio-spectral scene in a single two-dimensional measurement. It allows the capture of high-resolution spectral images using a single camera and a simple optical setup. Unlike traditional hyperspectral scanners that often capture images one pixel, row, or band

at a time, as in Figure 3.20, CASSI systems capture the spatial and spectral information contained in a scene in a single snapshot. The core principle of CASSI is to use a CA to modulate the incoming light as well as some optical shifting to distribute the energy from multiple wavelengths spatially. This encoded representation of an imaged scene is incident on a sensor which captures rich spatial and spectral information in a single snapshot. This encoded representation of an image can then be decoded using some reconstruction algorithm, such as the GPSR algorithm [165], which is commonly employed in the reconstruction of CASSI measurements [196, 212].

There are multiple CASSI architectures used within CS literature, each differing in how they encode spectral information or in their optical path compositions. These include but are not limited to Single Dispersion CASSI, Dual Dispersion CASSI, Colour CASSI, and 3D CASSI. Multi-shot sensing routines are common, in practice this requires using multiple unique CAs to capture multiple representations of the same scene in order to increase the information provided prior to reconstruction. Multi-shot CASSI has been shown to improve the quality of reconstructed CS measurement whilst trading the amount of compression realisable by the system. Thus, the design of ensembles of CAs which capture optimally distributed information and avoid redundancy becomes desirable.

As discussed in Section 2.5, an approach for performing TD directly on compressed measurements is desirable as it avoids; a) the need to capture a full representation of the scene using conventional spectral imaging, b) the potential errors and complexity associated with performing reconstruction. In this chapter, the fundamental operation of both SD-CASSI and DD-CASSI will be presented and the models used to simulate each system are derived. An overview of the proposed CASSI-based TD scheme is presented and finally the results of applying the CASSI-TD scheme to various multispectral datasets are shown.

6.2 Simulated CASSI Architectures

Two models of CASSI systems are considered in this chapter, these are SD-CASSI and DD-CASSI. Other models such as 3D CASSI [286] and architectures with side imaging [195, 207, 287] are not considered as these additions are generally to aid, or augment, in the reconstruction of the imaged scene which is not the focus of this work. Two simulation models were developed in MATLAB, each implementing discretised versions of either the single or dual dispersion CASSI architectures discussed in Section 3.7.2. These were developed to test the use of CASSI for TD and are idealised, *i.e.,* with no additive noise, and sensor, optical, and CA deposition defects are not modelled. The parameters used in these models are listed in Table 6.1.

Table 6.1: CASSI Simulation Parameters

Variable	Dimensions	Description
M		Number of columns in the hyperspectral datacube
N		Number of rows in the hyperspectral datacube
L		Number of bands in the hyperspectral datacube
K		Number of snapshots
F		Number of spectral filters
V		Resultant width of sheared measurement $(M + L - 1)$
$\mathbf{X}$	$M \times N \times L$	Hyperspectral datacube
$\boldsymbol{x}$	$MNL \times 1$	Vectorised hyperspectral datacube
$\boldsymbol{T}_{SD}$	$M \times N$	Coded aperture for single dispersion CASSI
$\boldsymbol{Y}_{SD}$	$V \times N$	Compressive measurement for single dispersion CASSI
$\boldsymbol{y}_{SD}$	$NV \times 1$	Vectorised compressive measurement for single dispersion CASSI
$\boldsymbol{H}_{SD}$	$NVK \times MNL$	Sensing matrix for single dispersion CASSI
$\boldsymbol{T}_{DD}$	$V \times N$	Coded aperture for dual dispersion CASSI
$\boldsymbol{Y}_{DD}$	$M \times N$	Compressive measurement for dual dispersion CASSI
$\boldsymbol{y}_{DD}$	$MN \times 1$	Vectorised compressive measurement for dual dispersion CASSI
$\boldsymbol{H}_{DD}$	$MNK \times MNL$	Sensing matrix for dual dispersion CASSI
$\boldsymbol{y}_x$	variable	Vectorised compressive measurement of an image
$\boldsymbol{y}_t$	variable	Vectorised compressive measurement of a target image
$\boldsymbol{y}_\Delta$	variable	Disparity between compressive image and target measurements
$\boldsymbol{Y}_x$	variable	Reshaped compressive measurement of an image
$\boldsymbol{Y}_t$	variable	Reshaped compressive measurement of a target image
$\boldsymbol{Y}_\Delta$	variable	Reshaped disparity between compressive image and target measurements

6.2.1 Single Dispersion CASSI

The SD-CASSI model, represented in Figure 3.29, can be discretised by assuming perfect transmission through optical elements and linear dispersion equal to the distance between pixels, Δd, between centre wavelengths of adjacent spectral channels. Based on this assumption, the optical flow through an SD-CASSI system, given in Equation (3.77), can be rewritten where the output of the detector is modelled as

$$\boldsymbol{Y}_{SD_{m,n}} = \sum_{l=1}^{L} \mathbf{X}_{(m-l+1),n,l} \boldsymbol{T}_{SD_{(m-l+1),n}} + \omega_{m,n} \tag{6.1}$$

where $m \in \{1, 2, ..., M+L-1\}$ and $n \in \{1, 2, ..., N\}$ are the spatial locations on the focal plane array, l denotes the spectral band $l \in L$, and ω is some residual noise term. The resultant discrete simulated optical flow through the model is shown in Figure 6.1.

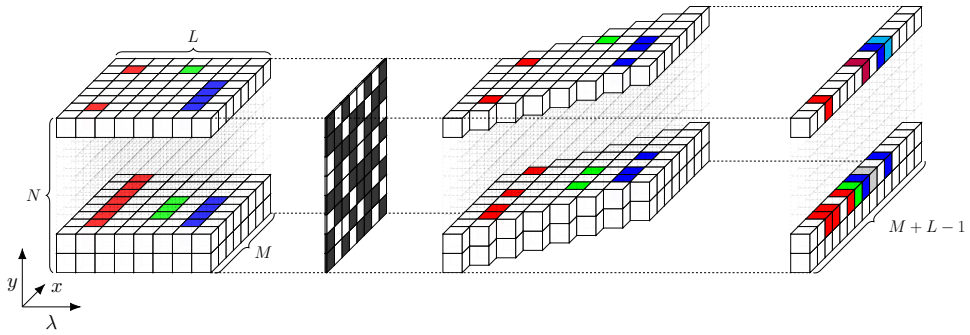


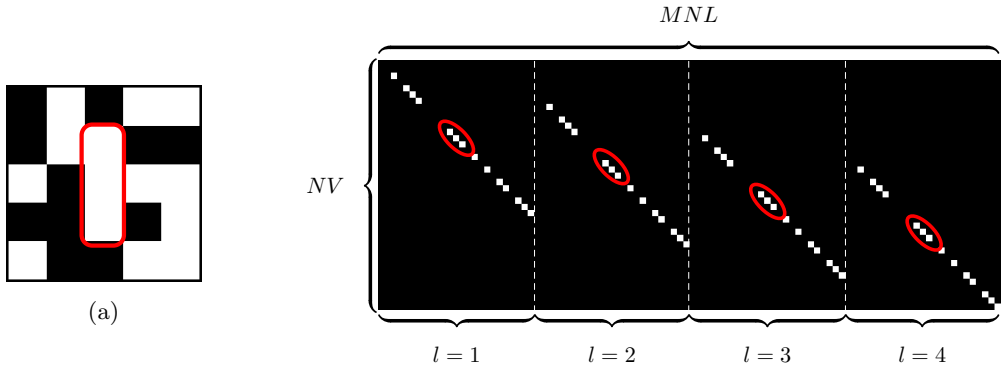
Figure 6.1: Simulated discrete optical flow through an SD-CASSI system.

Simplifying for discrete pixel locations and combining the shear term to form a sensing matrix allows Equation (6.1) to be written as a single matrix multiplication

$$\boldsymbol{y}_{SD} = \boldsymbol{H}_{SD}\boldsymbol{x} \tag{6.2}$$

where, $\boldsymbol{x}$, is some vectorised hyperspectral datacube, $\boldsymbol{y}$ is the observed CASSI measurement, and $\boldsymbol{H}_{SD}$ is the sensing matrix that describes the simulated encoding and shearing of the imaged scene passing through the SD-CASSI model. Equation (6.2) allows for a single CASSI measurement to be conducted by a single matrix multiplication.

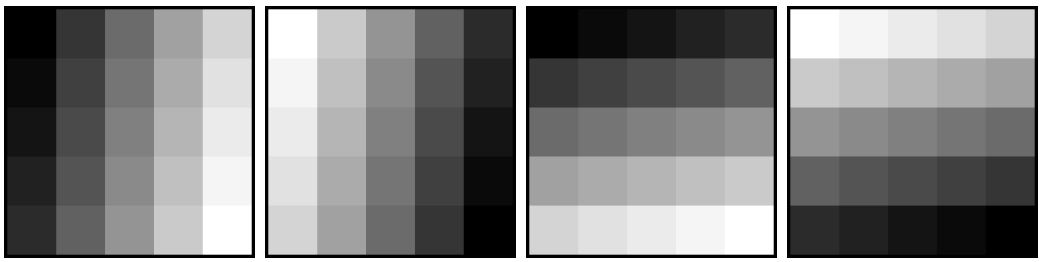
The sensing matrix, $\boldsymbol{H}_{SD}$, encodes the known bandpass and shifting behaviour for each pixel as it passes through the optical elements of the simulated SD-CASSI model. This behaviour consists of the encoding, determined by the CA, and a shear term, which is dependant on the wavelength of each band in a given pixel. Figure 6.2 shows a simple example of a Random CA, $\boldsymbol{T}_{SD}$, and the resulting sensing matrix, $\boldsymbol{H}_{SD}$, where $M = N = 5$, $L = 4$, and $K = 1$.



(a) (b)

Figure 6.2: Example of **a)** a Random binary CA and **b)** the associated sensing matrix for SD-CASSI.

The pixels highlighted in $\boldsymbol{T}_{SD}$, Figure 6.2a, appear four times in the associated sensing matrix $\boldsymbol{H}_{SD}$, Figure 6.2b. As $\boldsymbol{T}_{SD}$ is a binary bandpass/bandstop CA, these pixels affect the same elements in each of the L bands of the input but are shifted downwards in the sensing matrix, $\boldsymbol{H}_{SD}$. It is this shift that encodes the shearing operation simulating the prism element within the CASSI hardware, as spectral information from different spatial locations is coincident in the output. An example of how an image, Figure 6.3, passes through the simulated SD-CASSI hardware modelled by the sensing matrix in Figure 6.2b is shown in Figure 6.4.



(a) $l = 1$ (b) $l = 2$ (c) $l = 3$ (d) $l = 4$

Figure 6.3: Bands of an example synthetic multispectral image with $M = N = 5$, and $L = 4$.

The image from Figure 6.3 is first reshaped into a column vector, and multiplied by $\boldsymbol{H}_{SD}$ (as in Equation (6.2)) to form a single SD-CASSI measurement. The column vector is shown above the sensing matrix in Figure 6.4 in order to assist in visualising which pixel in the image is affected by each element in the sensing matrix, $\boldsymbol{H}_{SD}$.

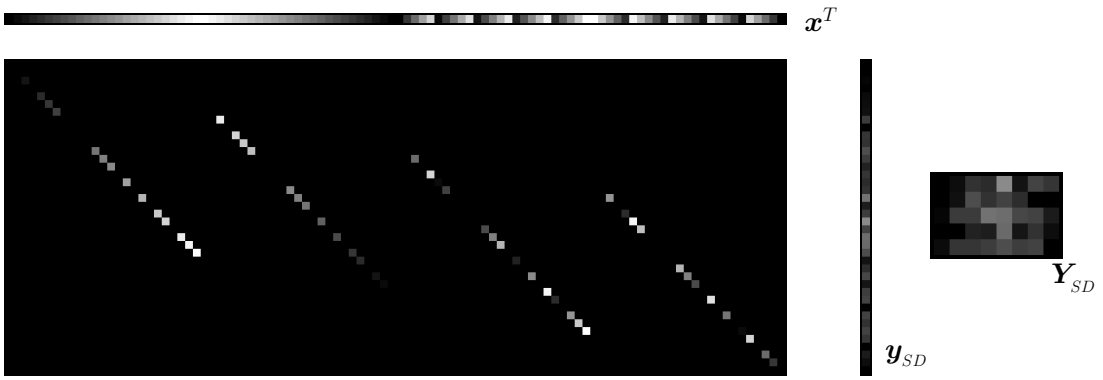


Figure 6.4: Example SD-CASSI measurement using the image from Figure 6.3.

From Figure 6.4 the operation of the simulated CASSI models is shown, for any pixel in a vectorised input image, $\boldsymbol{x}$, of size $1 \times MNL$, the columns of the sensing matrix, $\boldsymbol{H}_{SD}$, encode the new location of the energy contained in that pixel. The rows then indicate which of these elements are summed together to form the output.

For spatially rich and detailed scenes, a single CASSI measurement is often insufficient to wholly represent the information present [288]. The number of measurement snapshots can be increased in order to provide extra information,this is useful for both reconstruction and in the proposed TD scheme. During each snapshot, a different CA is used, allowing for varied information to be encoded in each k^{th} shot. In implementing a multi-shot SD-CASSI system, Equation (6.1) becomes

$$\boldsymbol{Y}^k_{SD_{m,n}} = \sum_{l=1}^{L} \mathbf{X}_{(m-l+1),n,l} \boldsymbol{T}^k_{(m-l+1),n} + \omega^k_{m,n} \quad k = 1, ..., K, \tag{6.3}$$

and the sensing matrix for each shot can be constructed as

$$\boldsymbol{y}^k_{SD} = \overbrace{\left[\begin{array}{c|c|c|c} \text{diag}(\boldsymbol{t}^k_{SD}) & \mathbf{0}_{M\times MN} & \cdots & \mathbf{0}_{M(L-1)\times MN} \\ & \text{diag}(\boldsymbol{t}^k_{SD}) & \cdots & \\ & & \ddots & \\ \mathbf{0}_{M(L-1)\times MN} & \mathbf{0}_{M(L-2)\times MN} & \cdots & \text{diag}(\boldsymbol{t}^k_{SD}) \end{array}\right]}^{\boldsymbol{H}^k_{SD}} \left[\begin{array}{c} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_L \end{array}\right] \tag{6.4}$$

The resultant multi-shot sensing matrix is constructed by stacking the individual sensing matrices corresponding to each shot, $\boldsymbol{H}^k_{SD}$. Figures 6.5 and 6.6 show simple examples of both Random and Boolean generated CAs and their associated sensing matrices, $\boldsymbol{H}_{SD}$, where $M = N = 5$, $L = 4$ and $K = 3$.

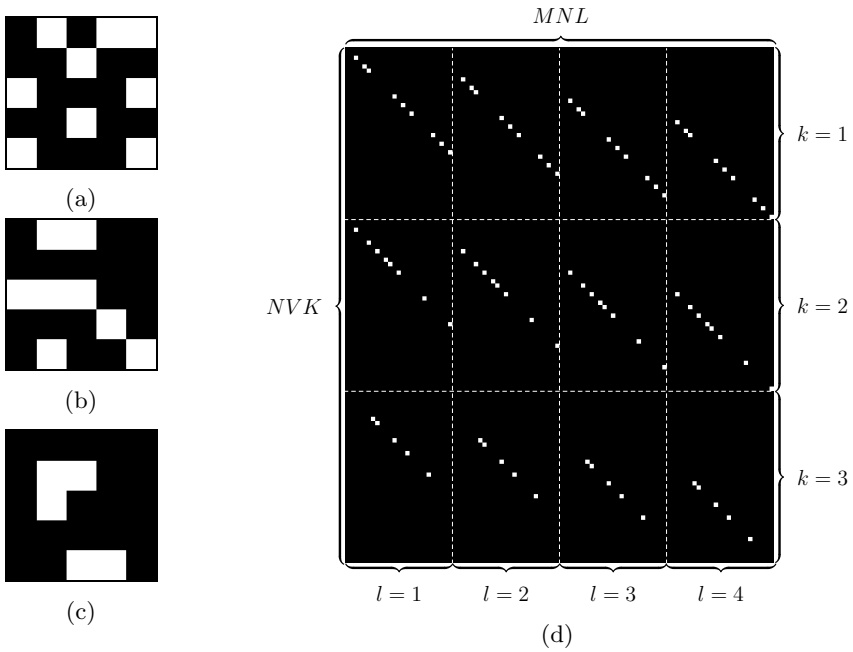


Figure 6.5: Examples of **a-c)** an ensemble of Random CAs and **d)** the resultant sensing matrix for multi-shot SD-CASSI.

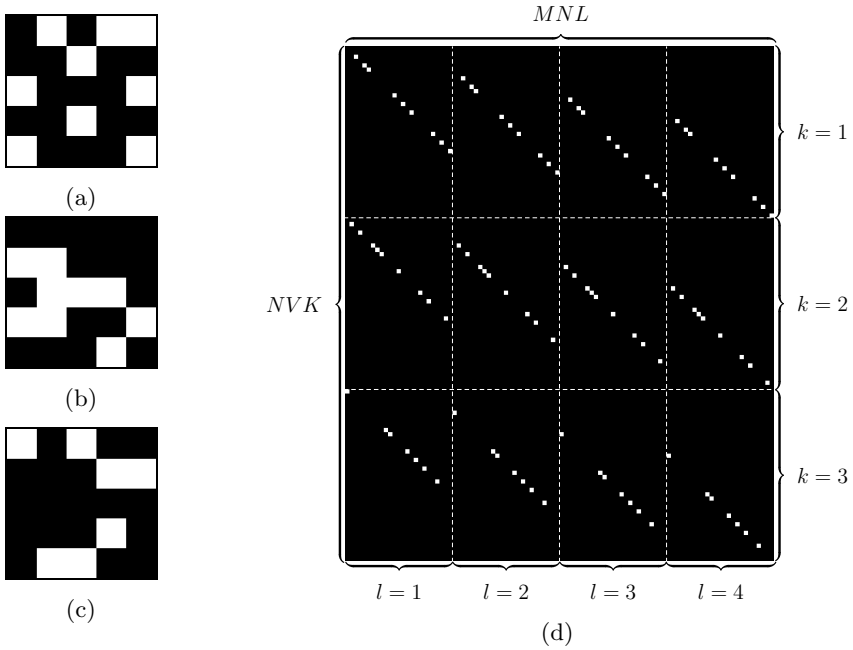


Figure 6.6: Examples of **a-c)** an ensemble of Boolean CAs and **d)** the resultant sensing matrix for multi-shot SD-CASSI.

6.2.2 Dual Dispersion CASSI

In addition to the SD-CASSI model, a dual dispersion, or DD-CASSI, model was created. Similarly to Equation (6.1) for a discretised SD-CASSI model, the equation for the optical flow through a DD-CASSI model (Equation (3.81)) can be discretised and rewritten as;

$$\boldsymbol{Y}_{DD_{m,n}} = \sum_{l=1}^{L} \mathbf{X}_{m,n,l} \boldsymbol{T}_{DD_{(m+l-1),n}} + \omega_{m,n}, \tag{6.5}$$

where $m \in \{1, 2, ..., M\}$ and $n \in \{1, 2, ..., N\}$ are the spatial locations on the focal plane array, l denotes the spectral band $l \in L$, and ω is some residual noise term. An example of this discretised DD-CASSI architecture is shown in Figure 6.7.

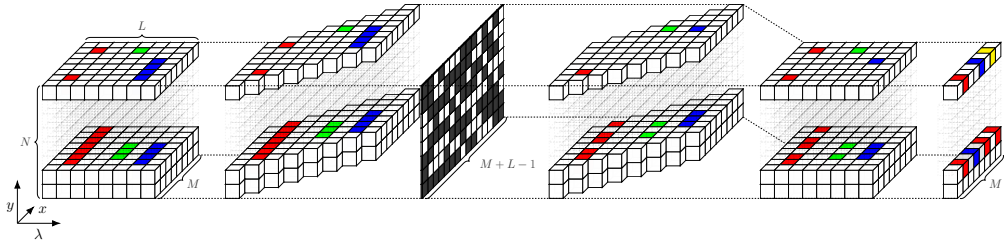


Figure 6.7: Simulated discrete optical flow through a DD-CASSI system.

Similarly to an SD-CASSI model, the DD-CASSI architecture described in Equation (6.5) can be formulated as a single matrix multiplication with a designed sensing matrix, $\boldsymbol{H}_{DD}$, as shown in Equation (6.6).

$$\boldsymbol{y}_{DD} = \boldsymbol{H}_{DD}\boldsymbol{x} \tag{6.6}$$

where, $\boldsymbol{x}$ is some vectorised hyperspectral datacube, $\boldsymbol{y}_{DD}$ is the observed DD-CASSI measurement, and $\boldsymbol{H}_{DD}$ is the sensing matrix that describes the simulated shearing, encoding, and reverse shearing of the imaged scene passing through the DD-CASSI model. As the input image is sheared prior to encoding, each band is passed through a different sub-region of the CA in the discretised model shown in Figure 6.7. An example of this is shown in Figure 6.8 for a Random CA, $\boldsymbol{T}_{DD}$, where $M = N = 5$ and $L = 4$, $\boldsymbol{T}_{DD}$ is of size $K \times N$ to account for the initial shearing operation.

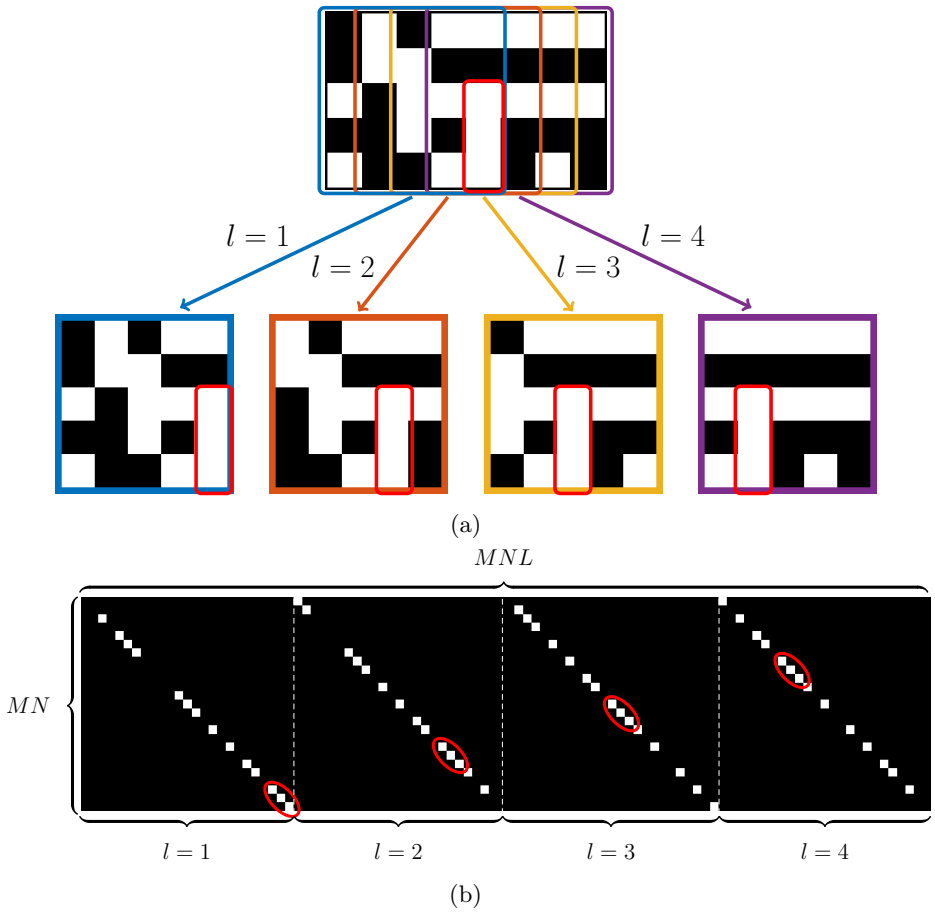


Figure 6.8: **a)** Example CA and **b)** corresponding sensing matrix for DD-CASSI.

Unlike in Figure 6.2, the sensing matrix, $\boldsymbol{H}_{DD}$, is created from different windows of $\boldsymbol{T}_{DD}$ as, following the initial shearing operation, each band is limited to a known spatial extent in the sheared image, as outlined in Figures 3.32b and 3.32c. Figure 6.8a shows how a CA, $\boldsymbol{T}_{DD}$, is used to simulate this behaviour in the sensing matrix, Figure 6.8b. The same region in the CA, highlighted by the red rectangle, corresponds to different spatial locations when encoding each band, l. Assigning different sub-regions of $\boldsymbol{T}_{DD}$ to each band emulates the initial dispersion prior to encoding, while the second, reverse, dispersion is implemented by removing the downwards shifting present in the example of SD-CASSI from Figure 6.2b. This results in a compressed measurement with the same spatial dimensionality as the original input.

An example of an image (Figure 6.3) passing through a simulated DD-CASSI architecture is shown in Figure 6.9. Similarly to the example for an SD-CASSI simulation, the vectors are placed to best represent which pixels are encoded by which elements in the sensing matrix, $\boldsymbol{H}_{DD}$, and how the output vector, $\boldsymbol{y}_{DD}$, is formed.

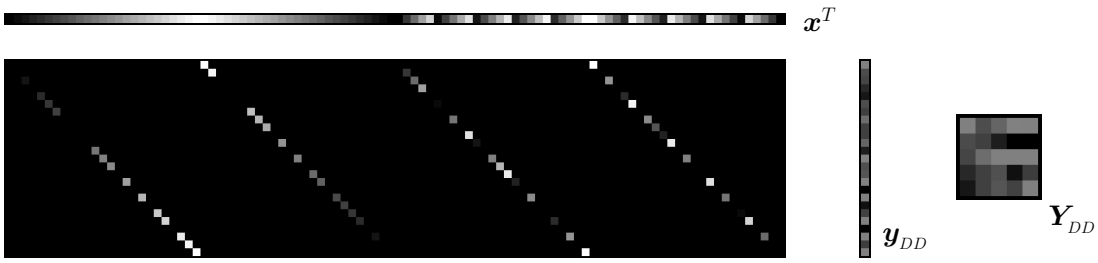


Figure 6.9: Example DD-CASSI measurement using the image from Figure 6.3.

Much like the multi-shot variant of an SD-CASSI model, a multi-shot DD-CASSI model can be achieved by taking repeated shots with a series of CAs, given in Equation (6.7),

$$\boldsymbol{Y}^k_{DD_{n,m}} = \sum_{l=1}^{L} \mathbf{X}_{n,m,l} \boldsymbol{T}^k_{n,(m+l-1)} + \omega^k_{n,m} \quad k = 1, ..., K. \tag{6.7}$$

Similarly to the example for constructing the k^{th} shot in an SD-CASSI system in Equation (6.4), the k^{th} shot of a DD-CASSI system can be written in matrix form where the resultant multi-shot sensing matrix, $\boldsymbol{H}_{DD}$, can be created by stacking multiple individual sensing matrices associated with each CA,

$$\boldsymbol{y}^k_{DD} = \overbrace{\left[\text{ diag}\left(\boldsymbol{t}^k_{DD}\right)|_{l=1} \;\middle|\; \text{diag}\left(\boldsymbol{t}^k_{DD}\right)|_{l=2} \;\middle|\; \cdots \;\middle|\; \text{diag}\left(\boldsymbol{t}^k_{DD}\right)|_{l=L} \right]}^{\boldsymbol{H}^k_{DD}} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_L \end{bmatrix} \tag{6.8}$$

Examples of multi-shot DD-CASSI systems are shown in Figures 6.10 and 6.11 for ensembles of Random and Boolean CAs respectively, where $M = N = 5$, $L = 4$, and $K = 3$.

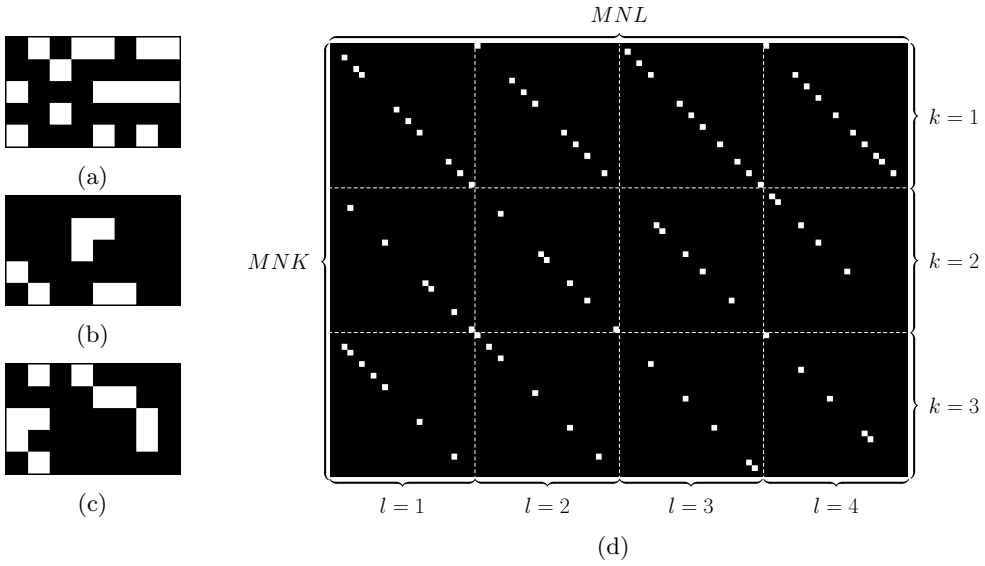


Figure 6.10: Examples of **a-c)** an ensemble of Random CAs and **d)** the resultant sensing matrix for multi-shot DD-CASSI.

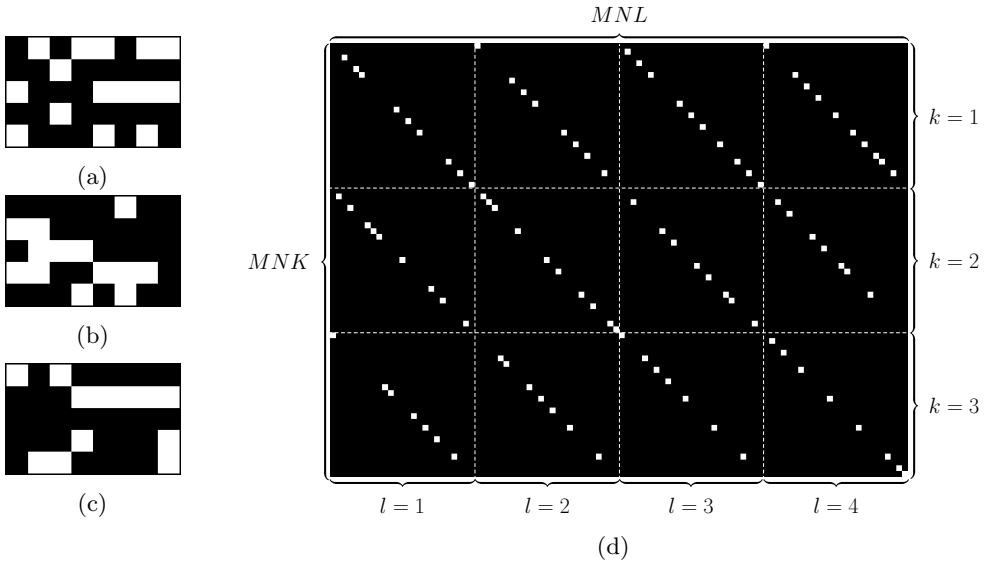
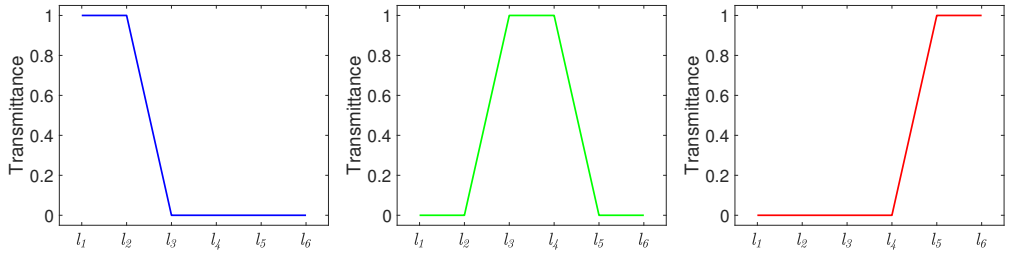


Figure 6.11: Examples of **a-c)** an ensemble of Boolean CAs and **d)** the resultant sensing matrix for multi-shot DD-CASSI.

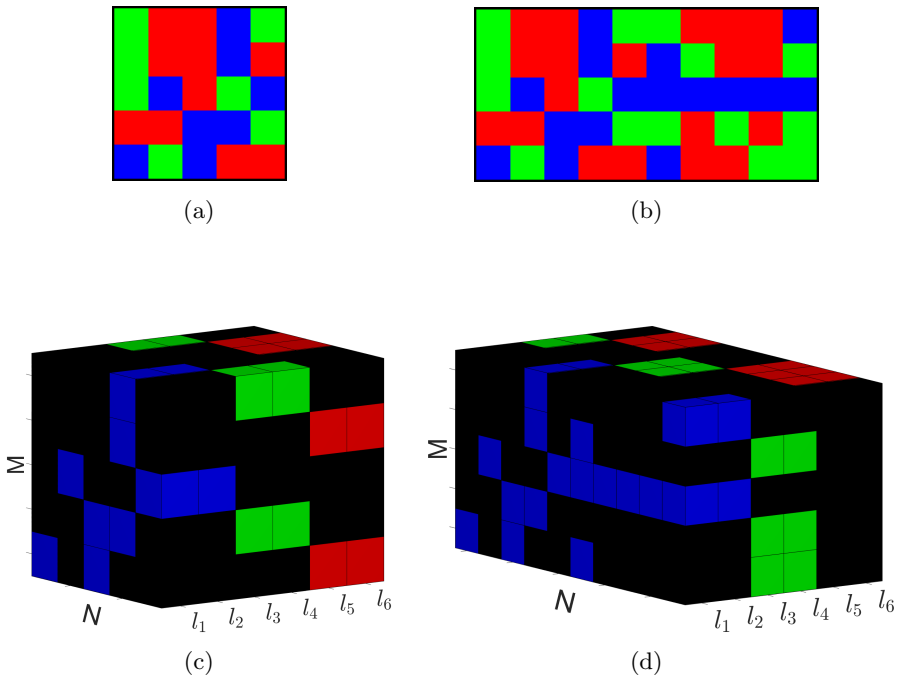
6.2.3 Colour CASSI Simulation

The definition of Colour Coded Aperture Snapshot Spectral Imaging (CCASSI) is near identical to its binary counterpart. However, rather than including bandpass/bandstop CAs over the range of spectral bands considered by the CASSI model, each pixel in a colour CA takes on a predefined response based on a set of desired spectral filters, F. An example set of filters, where $L = 6$, is shown in Figure 6.12 and a pair of Boolean CAs created using them are shown in Figure 6.13. Unlike in Figure 3.35, where colour indicated which shot in an ensemble CA the pixel was sampled in, the colour of each pixel indicates which filter is used.



(a) (b) (c)

Figure 6.12: Example filters used in Colour CA generation.



(a) (b) (c) (d)

Figure 6.13: Example colour CAs created using the filters from Figure 6.12. **a)** Boolean single dispersion colour CA. **b)** Boolean dual dispersion colour CA. **c)** Volumetric representation of Figure 6.13a. **d)** Volumetric representation of Figure 6.13b.

6.2.3.1 Single Dispersion CCASSI

The equation for forming a Single Dispersion Colour CASSI (SD-CCASSI) sensing matrix is nearly identical to that of its binary equivalent in Equation (6.4) with an additional spectral constraint, *i.e.,* the diagonal encoding components become $\text{diag}\,(\boldsymbol{t}^k_{SD})|_{l=1}$ in place of $\text{diag}(\boldsymbol{t}^k_{SD})$. This additional constraint ensures only CA elements corresponding to the correct filters are included in the final sensing matrix. With this additional constraint included, the SD-CCASSI sensing matrix for single, or multiple, shots can be derived as

$$\boldsymbol{y}^k_{DD} = \overbrace{\begin{bmatrix} \text{diag}\,(\boldsymbol{t}^k_{SD})|_{l=1} & \boldsymbol{0}_{M\times N^2} & \cdots & \boldsymbol{0}_{M(L-1)\times N^2} \\ & \text{diag}\,(\boldsymbol{t}^k_{SD})|_{l=2} & \cdots & \\ & & \ddots & \\ \boldsymbol{0}_{M(L-1)\times N^2} & \boldsymbol{0}_{M(L-2)\times N^2} & \cdots & \text{diag}\,(\boldsymbol{t}^k_{SD})|_{l=L} \end{bmatrix}}^{\boldsymbol{H}^k_{SD}} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_L \end{bmatrix} \tag{6.9}$$

If binary coded apertures are used, then $\boldsymbol{t}_{SD_1^k} = \boldsymbol{t}_{SD_2^k} = \boldsymbol{t}_{SD_L^k}$ and Equation (6.9) becomes equivalent to Equation (6.4). An example SD-CCASSI sensing matrix derived from a Random colour CA where $M = N = 5$, $L = 6$, and $F = 3$ is shown in Figure 6.14. Multi-shot SD-CCASSI sensing matrices can also be generated in the same way as other architectures. An example of multi-shot SD-CCASSI sensing matrices based on Random and Boolean CAs are shown in Figures 6.15 and 6.16 respectively.

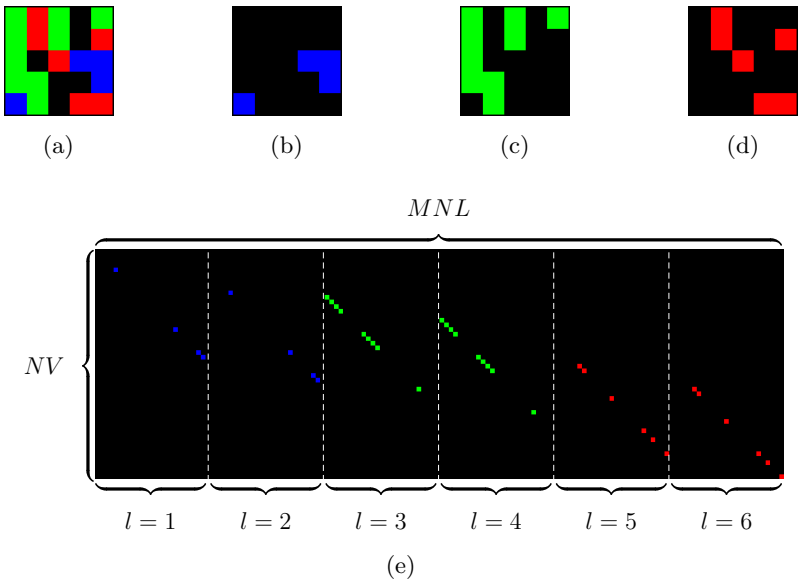


Figure 6.14: **a)** Random colour CA. **b-d)** Individual colour filter components of Figure 6.14a. **e)** Sensing matrix for SD-CCASSI.

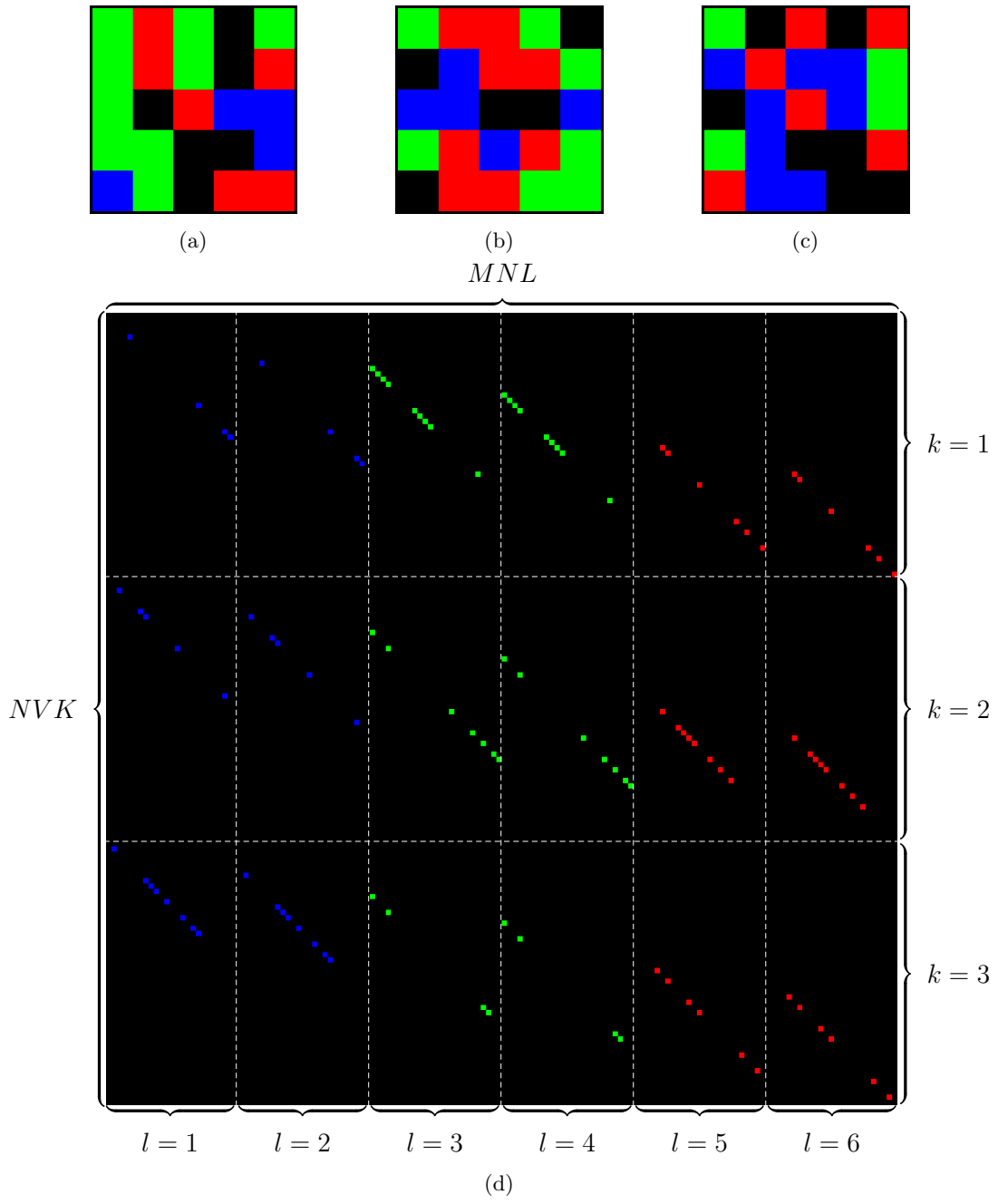


Figure 6.15: Examples of **a-c)** an ensemble of Random CAs and **d)** the resultant sensing matrix for SD-CCASSI

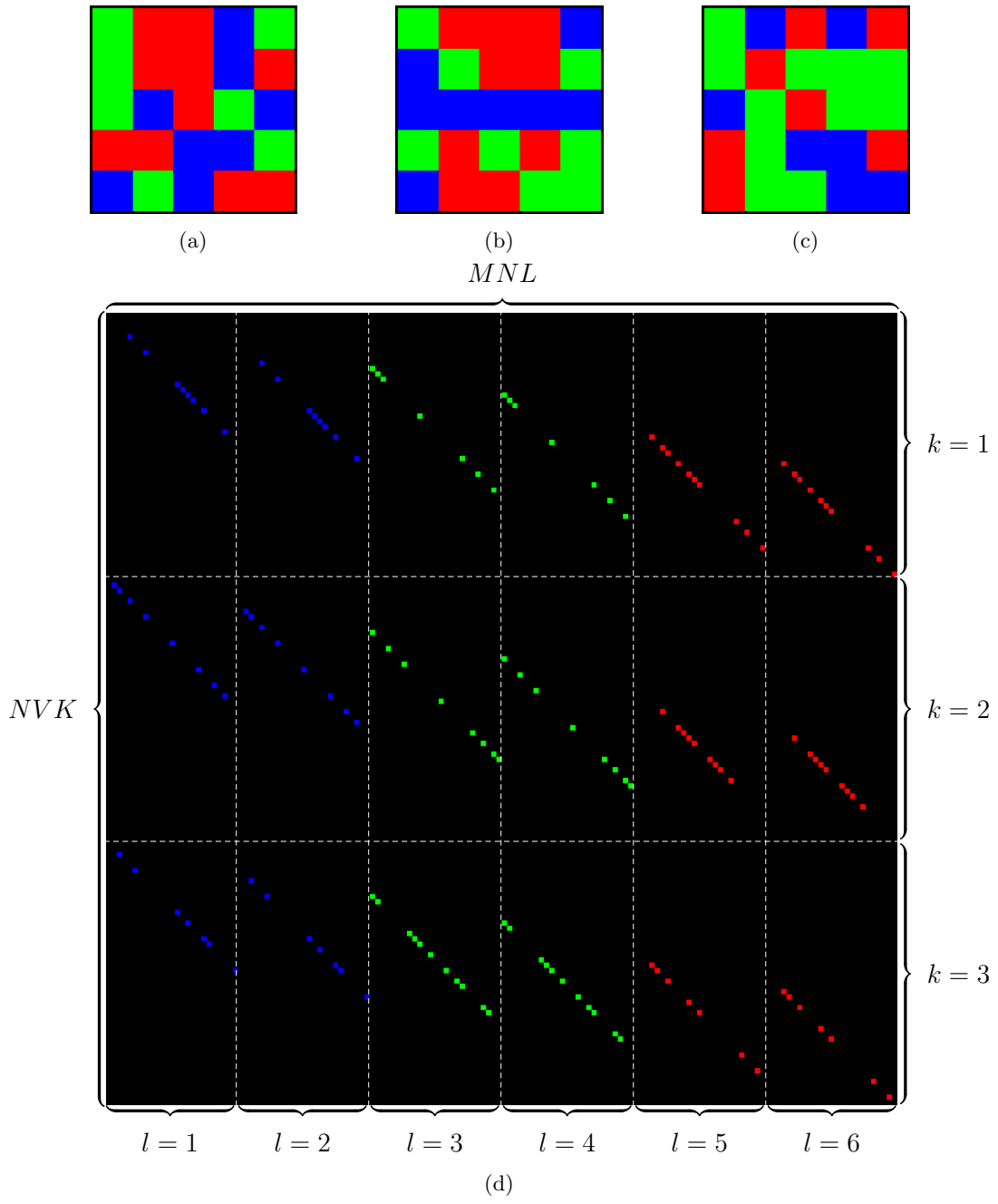


Figure 6.16: Examples of **a-c)** an ensemble of Boolean CAs and **d)** the resultant sensing matrix for SD-CCASSI.

6.2.3.2 Dual Dispersion CCASSI

The final CASSI architecture considered in this thesis is Dual Dispersion Colour CASSI (DD-CCASSI). The derivation of the DD-CCASSI sensing matrix is identical to that shown in Equation (6.8) for DD-CASSI with the additional constraint that the band dependant elements in $\text{diag}\,(\boldsymbol{t}^k_{DD})|_l$ also account for the response of filter, f, at that band, l.

$$\boldsymbol{y}^k_{DD} = \overbrace{\left[\ \text{diag}\,(\boldsymbol{t}^k_{DD})|_{l=1} \ \middle|\ \text{diag}\,(\boldsymbol{t}^k_{DD})|_{l=2} \ \middle|\ \cdots \ \middle|\ \text{diag}\,(\boldsymbol{t}^k_{DD})|_{l=L}\ \right]}^{\boldsymbol{H}^k_{DD}} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_L \end{bmatrix} \tag{6.10}$$

An example DD-CCASSI sensing matrix derived from a Random colour CA where $M = N = 5$, $L = 6$, and $F = 3$ is shown in Figure 6.17. Similarly to the binary case (Figure 6.8), each band in an input image is encoded with a different region of the colour CA. As in SD-CCASSI (Figure 6.14), this encoding is further constrained, compared to the binary case, to locations where the filter element in the colour CA acts as a bandpass. Multi-shot DD-CCASSI sensing matrices, where $K = 3$, derived from Random and Boolean colour CAs are shown in Figures 6.18 and 6.19 respectively.

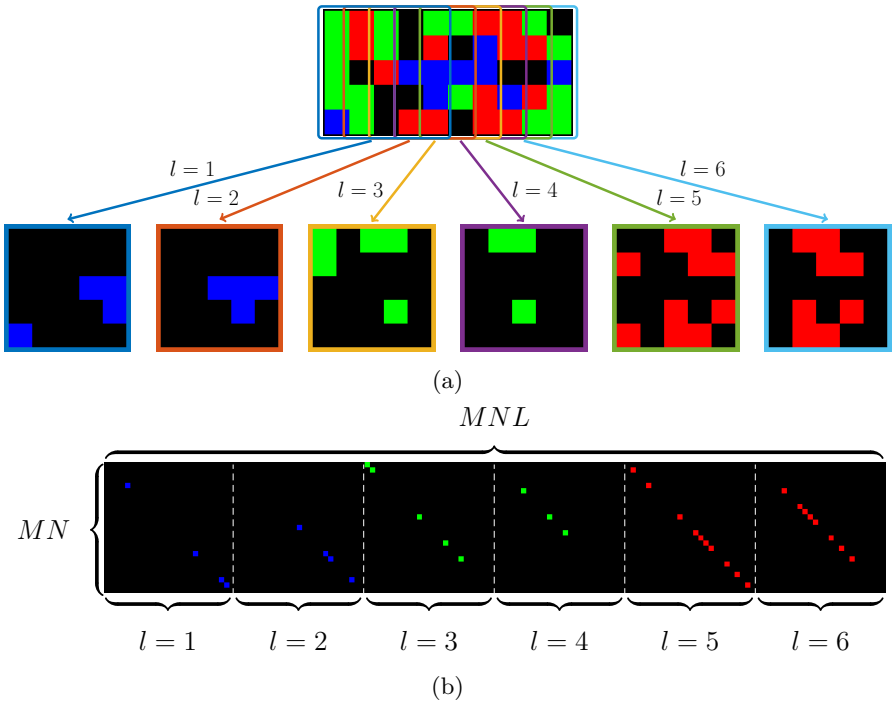


Figure 6.17: **a)** Example CA and **b)** corresponding sensing matrix for DD-CCASSI.

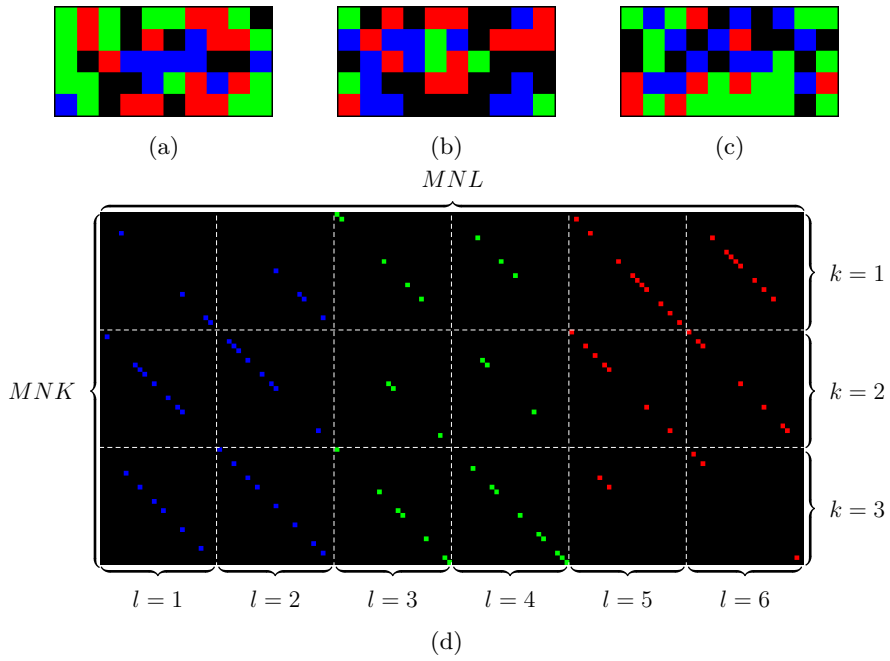


Figure 6.18: Examples of **a-c)** an ensemble of Random colour CAs and **d)** the resultant sensing matrix for DD-CCASSI.

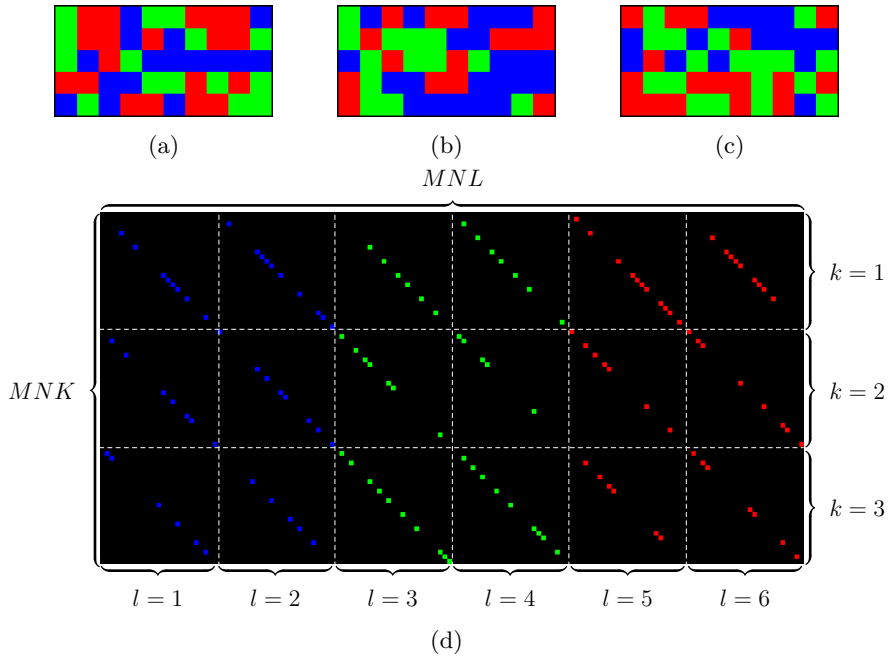


Figure 6.19: Examples of **a-c)** an ensemble of Boolean colour CAs and **d)** the resultant sensing matrix for DD-CCASSI.

6.3 Target Detection using a Simulated Coded Aperture Snapshot Spectral Camera

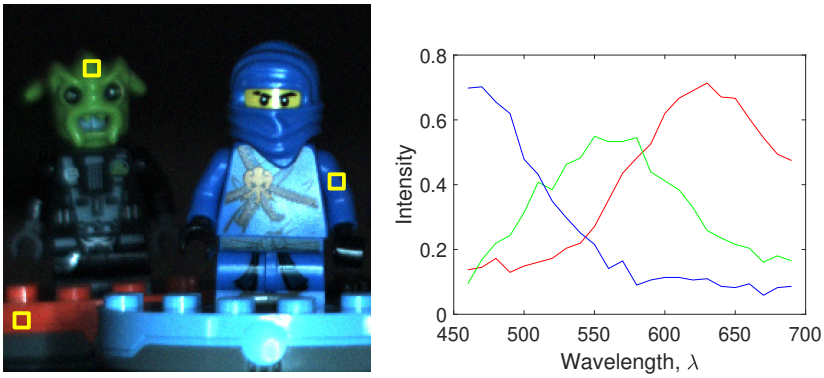
To date, pixel-wise object detection has not been explored extensively in a projected CS domain, with limited studies on machine learning-based classification from CASSI [289] and CS-MUSI [209] measurements. Instead object detection and tracking in CS research commonly exploit sparsity in order to separate object from background [171, 290] prior to any further processing or through the use of *a-priori* learned features [283]. The ability to detect and classify on a pixel-wise level is important in long range and aerial object detection applications as there may only be a handful of pixels on target. This section discusses the design of a novel TD framework using CASSI architectures, CASSI-TD.

Initially, the SD-CASSI model was simulated and tested to investigate the effectiveness of a TD approach, however, due to the shearing effect, the energy of targets was spread across multiple pixels and performance suffered as a result, especially around the border of the imaged scene and on the edges of objects. It can be observed that three approaches can be taken in response to this; 1) to continue with the current scenario where boundary information may be lost from detected objects due to the smearing effects from neighbouring pixels, 2) to create a detection algorithm that searches over multiple pixels for the dispersed energy of the target, and 3) to instead utilise a DD-CASSI system which, although more complex and requiring finer calibration when using a physical imaging system, allows for the structure of the input data to be preserved. Options 1 and 3 are explored in this thesis to compare how the two models can be applied to the TD problem.

As discussed in Section 2.5, CASSI architectures are heavily dependent on the design of CAs for any particular application. The use of both binary and colour CAs along with various design schemes and, in the case of CCASSI, the use of spectral selectivity allows for investigation into the suitability of the CASSI architecture for application to TD problems. This is desirable as it would result in 1) the ability to detect targets in a captured measurement without the need to capture the whole image using standard imaging technologies prior to compression via transform coding or DR and 2) avoid the need to reconstruct these compressed measurements prior to performing TD on the estimated image.

6.3.1 Target Detection in the Compressed Domain

With the development of models for each of the architectures considered, methods for performing TD using compressive CASSI measurements were investigated. A test image, provided by the University of Delaware's Computational Imaging and Machine Learning Laboratory, is shown in Figure 6.20.



(a) (b)

Figure 6.20: **a)** Example Multispectral Image used in CASSI-TD with annotated target material ground truth locations, and **b)** resultant ground truth spectra.

Three target objects were selected from the image; the red brick, the green head, and the blue outfit. These targets were selected as they are distinct from the background and therefore have a high probability of detection in the sparse domain. The regions selected as ground truth target spectra are highlighted in Figure 6.20a and the resultant average target spectrum for each object is shown in Figure 6.20b.

The behaviour for each pixel in an image, $\mathbf{X}$, with spatial dimensions M and N, and L spectral bands, as it passes through the optics of a CASSI system is known. This is due to the structure of the sensing matrix, $\boldsymbol{H}$, being known, given some dispersion behaviour and a designed CA. This means that the behaviours at locations in the image where there are objects of interest is also predictable. The output of any of the discussed CASSI architectures can be formulated as

$$\boldsymbol{y}_x = \boldsymbol{H}\boldsymbol{x} \tag{6.11}$$

where $\boldsymbol{x} \in \mathbb{R}^{1\times MNL}$ is the vectorised version of image, $\mathbf{X}$, and $\boldsymbol{y}_x$ is the compressed measurement given some sensing matrix, $\boldsymbol{H}$.

Given a reference spectrum of interest, $\boldsymbol{r} \in \mathbb{R}^{1\times L}$, a synthetic target image, $\mathbf{T}$, can be constructed by replicating the reference spectrum over the spatial extent of the image under

test. The CASSI measurement of such a target reference image is calculated using the same sensing matrix, $\boldsymbol{H}$, as

$$\boldsymbol{y}_t = \boldsymbol{H}\boldsymbol{t}, \tag{6.12}$$

where $\boldsymbol{t} \in \mathbb{R}^{1\times MNL}$ is the vectorised version of image, $\mathbf{T}$, and $\boldsymbol{y}_t$ is the compressed measurement of the target given that same sensing matrix, $\boldsymbol{H}$. This process is identical for both CASSI and CCASSI, with the different behaviours described using variations in the sensing matrices.

As the behaviour through the optical elements is encoded by the sensing matrix, $\boldsymbol{H}$, comparisons between $\mathbf{X}$ and $\mathbf{T}$ can be performed using their respective compressed measurements. In a DD-CASSI system, if some pixel in both $\mathbf{X}$ and $\mathbf{T}$ are identical, $\boldsymbol{Y}_x$ and $\boldsymbol{Y}_t$ will also be identical as the pixel in the measurement is dependent only on the corresponding pixel in the observed image, as given in Equation (6.5). However, in an SD-CASSI system, the compressed measurement at some pixel is influenced by the spectra of adjacent pixels, as derived in Equation (6.1). Unless there are sufficiently large ($L-1$ wide) regions of contiguous target material, non-target materials may have an overwhelming influence over any target point and the resultant compressed representation in any SD-CASSI measurement.

Comparing $\boldsymbol{Y}_x$ and $\boldsymbol{Y}_t$ and identifying regions where the difference between them is low forms the basis of the proposed CASSI-TD approach. The simplest form of comparison is the absolute difference between the two measurements

$$\boldsymbol{y}_\Delta = |\boldsymbol{y}_x - \boldsymbol{y}_t| \tag{6.13}$$

where, areas of agreement between $\boldsymbol{y}_x$ and $\boldsymbol{y}_t$ will result in low values in $\boldsymbol{y}_\Delta$. $\boldsymbol{y}_\Delta$ can be reshaped into a measurement image, $\boldsymbol{Y}_\Delta$, which can be thresholded in order to produce a binary detection image, containing the set of pixels predicted as the target class. Other, more robust or informative comparisons (such as those incorporating spatial information) may be applied, however, the absolute difference is used here as an example. The CASSI-TD approach is shown visually in Figure 6.21. Additionally, detections based on each of the target spectra presented in Figure 6.20b are obtained and displayed in Figure 6.22.

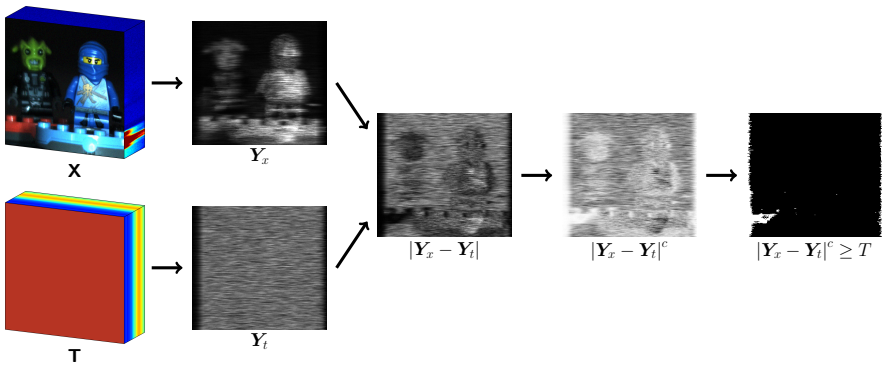


Figure 6.21: Example of the CASSI-TD pipeline applied to the example image shown in Figure 6.20.

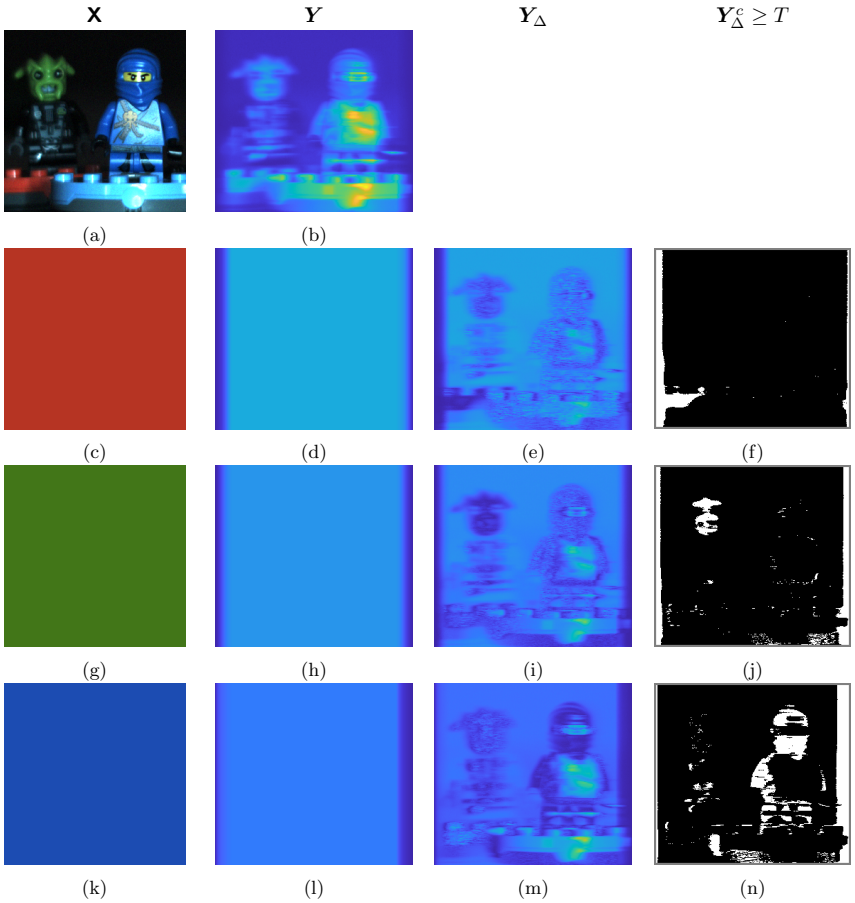


Figure 6.22: Example detections using a CASSI-TD architecture. **a)** Pseudo-colour RGB image. **b)** SD-CASSI representation of Figure 6.22a with $K = 10$. **c, g, k)** Target images based on the spectra from Figure 6.20b. **d, h, l)** SD-CASSI representation of Figures 6.22c, 6.22g and 6.22k with $K = 10$. **e, i, m)** Absolute difference between Figures 6.22d, 6.22h and 6.22l andFigure 6.22b. **f, j, n)** Binary detections from Figures 6.22e, 6.22i and 6.22m with $T = 0.9$.

6.4 Results

In this section, the results of testing the simulated CASSI architectures on multiple MSI datasets of varying complexity to determine their efficacy in TD applications are presented. First, a benchmark test using a synthetic colour checker is performed to assess how each of the architectures compare to one another. Additionally, tests on bench-top imagery from the CAVE dataset [291] (Figure A.2.1) are carried out to ascertain how CASSI-TD performs on similar images collected in a controlled lab environment. Finally, images from the UDRC dataset (Figure A.3.2) are spectrally downsampled to be comparable to the CAVE imagery in order to test the performance of CASSI-TD in a remote sensing application with relatively small targets.

6.4.1 CASSI-TD Results on Synthetic Multispectral Imagery

6.4.1.1 Data Description

In order to benchmark each of the simulated CASSI architectures against one another, as well as gain an initial insight into how they perform in a TD application, a synthetic image based on a colour checker, a common tool used in colour correction, was created. The image consisted of twenty-four target objects with varied spectral profiles. Each of the targets is intended to mimic the spectral response of a common material or colour [292]. The image dimensions are as follows; $M = 440$, $N = 296$, and $L = 36$. Each individual target is an $M = N = 64$ square of pixels with an 8 pixel buffer zone between each square. The spectral responses are based on measurements taken by the Munsell Color Science Lab at Rochester Institute of Technology [293]. A pseudo-colour RGB image of the synthetic dataset is shown in Figure 6.23 with each of the twenty-four targets' spectral responses in the range considered graphed in Figure 6.24.

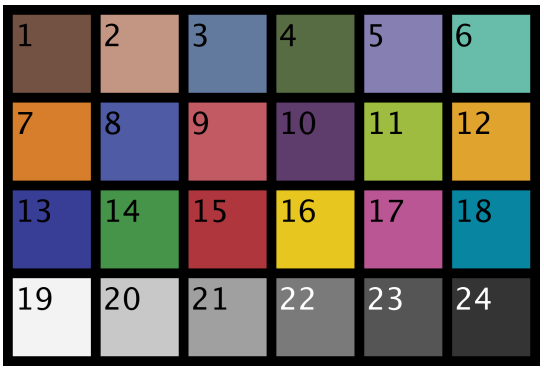


Figure 6.23: Synthetic Gretag-Macbeth Colour Checker used to test the CASSI-TD simulations with overlaid target numbers.

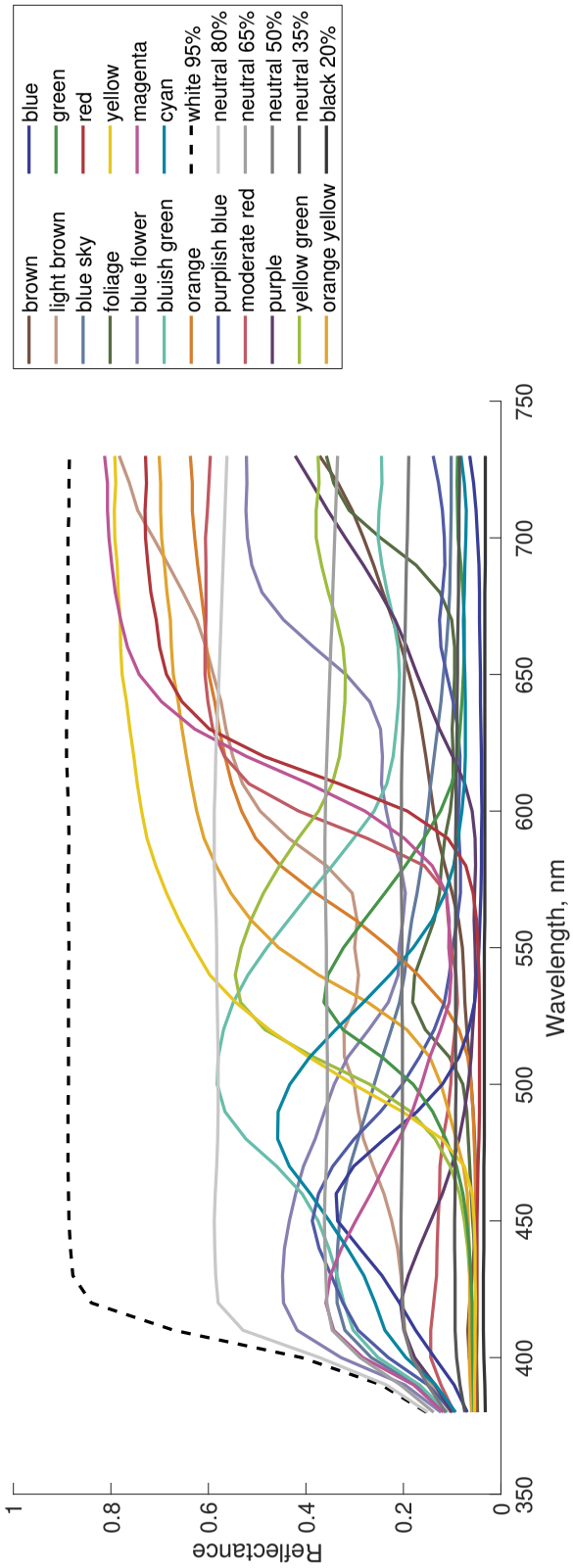
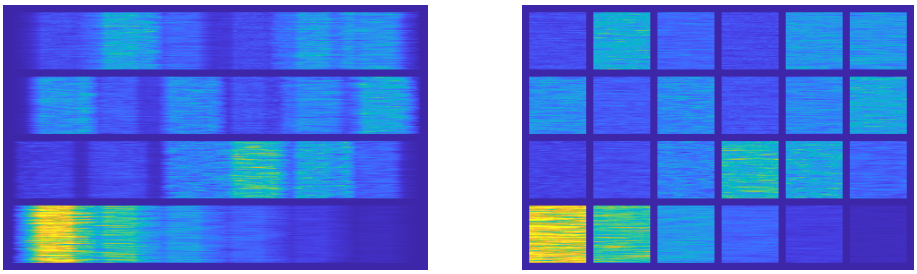


Figure 6.24: Target Spectra present in Gretag-Macbeth Colour Checker [293]

6.4.1.2 Binary CASSI-TD Performance on Synthetic Macbeth Chart

In order to assess the efficacy of detecting target spectra from compressed CASSI measurements, each image and series of corresponding target images were passed through each of the simulated CASSI architectures (as pictured in Figure 6.21). By varying the type of CAs and number of shots, K, used to generate the sensing matrix, $\boldsymbol{H}$, a number of compressed representations of the synthetic colour checker image (Figure 6.20) were generated. Example of such compressed measurements are shown in Figure 6.25 for both SD-CASSI and DD-CASSI.



(a) (b)

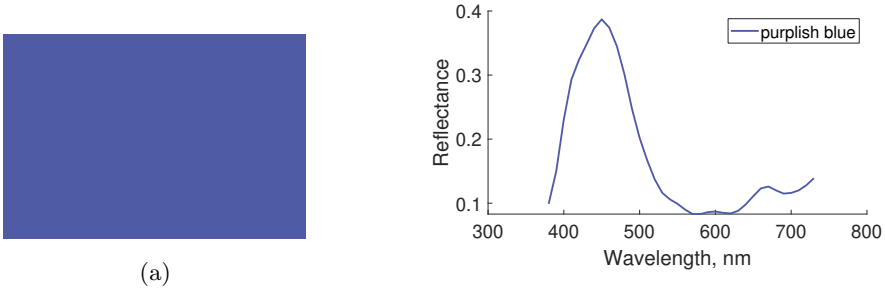
Figure 6.25: Compressed measurements of Figure 6.23 from **a)** SD-CASSI, and **b)** DD-CASSI, using a Random CA with $K = 10$ shots.

Figure 6.25 shows the compressed measurements based on a Random CA with $K = 10$ shots from SD-CASSI and DD-CASSI. The dark blue colours indicate a low energy measurement, or dark pixel at the output, whereas the yellow colours indicate a high energy measurement, or a bright pixel at the output. From Figures 6.25a and 6.25b it is apparent that bright targets such as the white and yellow tiles (targets 19 and 16 respectively in this case) are correspondingly bright in the compressed measurement. The opposite also holds true, with the black, blue, and purple tiles (targets 24, 13, and 10 respectively) having a relatively low brightness. However, this energy is not distributed uniformly within the extent of each target due to the random sampling produced by the CA. This is clearly visible in the examples shown in Figure 6.25, however, is particularly noticeable in Figure 6.25a where the energy of an image is more distributed due to the single shearing operation in SD-CASSI.

Additionally, there is limited ability to distinguish between spectra in a binary CASSI system where two spectra have similar AUC (Figure 6.24), especially in the DD-CASSI measurement, as any the compressed observation is largely equivalent to the energy present in an input pixel (with some variation on account of the encoding being applied to the sheared spectra). Therefore, if two spectra have similar energy in the spectral domain, they will have similar compressive observations in a DD-CASSI measurement. In the SD-CASSI case, there is slightly more variation in the intensities observed, but this can be attributed to the mixing of spectra

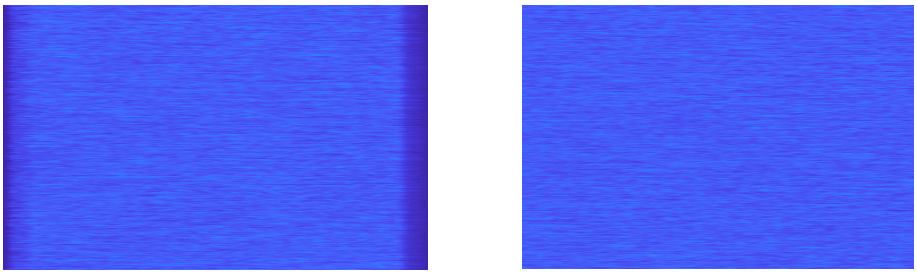
in each pixel of the compressed measurement. As there is no reverse shearing operation, pixels in an SD-CASSI measurement are mixtures of up to $L-1$ adjacent pixels, with exceptions at the boundary of a measurement (as displayed in Figure 6.1), where fewer pixels contribute to the integration on the focal plane array. This causes the gradual darkening on the left and right of Figure 6.25a.

Each of the target spectra (Figure 6.24) were used to create target images, $\mathbf{T}$, which were passed through the same sensing matrices as the image scene. An example of such a target image for target 8 ("purplish blue") is shown in Figure 6.26a where this has the same dimensionality as the image it will be compared against, $\mathbf{X}$. The spectrum used to create the target image in Figure 6.26a is shown in Figure 6.26b, where every N-dimensional pixel $\boldsymbol{t} \in \mathbf{T}$ corresponds to this spectrum. The compressed measurements for the target image, $\mathbf{T}$, captured using SD-CASSI and DD-CASSI are displayed in Figure 6.27.



(a) (b)

Figure 6.26: Example target Image used in CASSI-TD. **a)** Image composed of target 8 "purplish blue". **b)** Spectrum of target 8 used to create a).



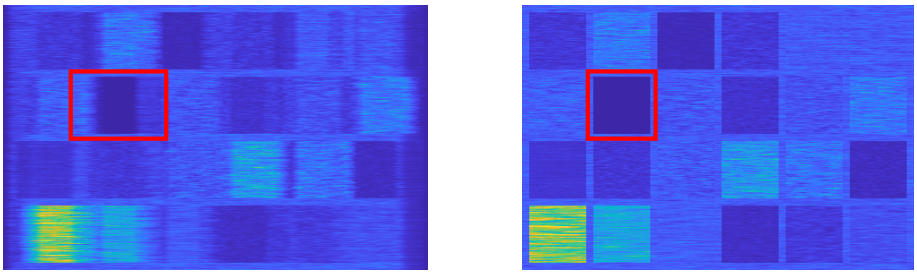
(a) (b)

Figure 6.27: Compressed measurements of the "purplish blue" image from Figure 6.26a captured using **a)** SD-CASSI, **b)** DD-CASSI, using identical sensing matrices to those in Figure 6.25.

Both the image and target measurements alone have little predictive value from a TD perspective; with the compressed image measurement containing similar values for different spectra, and the target image measurement being highly monotonous. However, as the

behaviour for any two identical co-located pixels in both of these measurements created by the same sensing matrix should be indistinguishable, comparing these measurements can locate regions of interest.

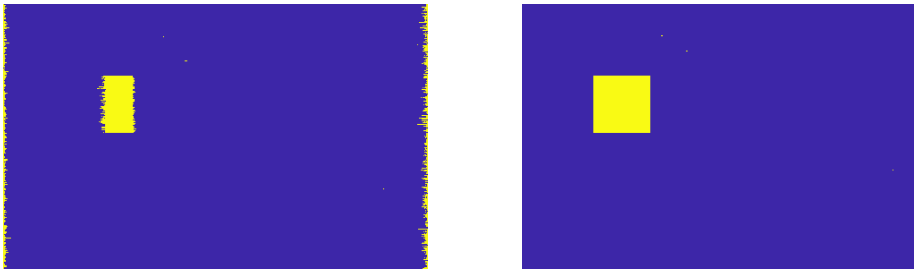
The results of taking the absolute difference between Figures 6.25a and 6.27a, and Figures 6.25b and 6.27b are shown in Figure 6.28 for SD-CASSI and DD-CASSI respectively, where dark blue colours indicate small differences and yellow colours indicate larger differences.



(a) (b)

Figure 6.28: Absolute differences between image (Figure 6.25) and target (Figure 6.27) measurements using **a)** SD-CASSI, and **b)** DD-CASSI.

From Figure 6.28, both the SD-CASSI and DD-CASSI-based TD methods isolate the target of interest reasonably well, with the tile corresponding to the "purplish blue" material being visibly darker than the surroundings. Most similar in either case, is the location corresponding to "blue sky" (target 3). By thresholding the complement of this result at a series of thresholds, $T \in \{0, 0.1, ..., 0.99, 1\}$, a set of binary classification metrics at each threshold to quantify each method's performance can be obtained. An example of this thresholding operation, at $T = 0.99$ is shown in Figure 6.29a and Figure 6.29b for SD-CASSI and DD-CASSI respectively.



(a) (b)

Figure 6.29: CASSI-TD predictions of the "purplish blue" target in Figure 6.24 using a Random CA with $T = 0.99$ using **a)** SD-CASSI and, **b)** DD-CASSI.

Figure 6.29 shows the results of applying a threshold to the difference between the image and target measurements, the SD-CASSI method achieves an F1 score of 0.4357, whereas

the DD-CASSI method achieved a score of 0.9994. As discussed throughout this chapter, as each pixel in the SD-CASSI measurement is comprised of components from multiple pixels in the source image, the ability to detect targets using this technique is limited as the image measurement will appear different from the expected target response as the target image is comprised of pure spectra only. Alternative algorithms that take into account this known spatial distribution of spectral information may improve the performance of any SD-CASSI-based TD techniques.

The compressed measurements of Figure 6.23, absolute differences to some target image, and thresholded results using a Boolean CA in conjunction with both a SD-CASSI and DD-CASSI architecture are shown in Figure 6.30. Using a Boolean CA produces a final F1 score at $T = 0.99$ of 0.4452 for SD-CASSI and 1 for DD-CASSI, a slight improvement over the Random CAs.

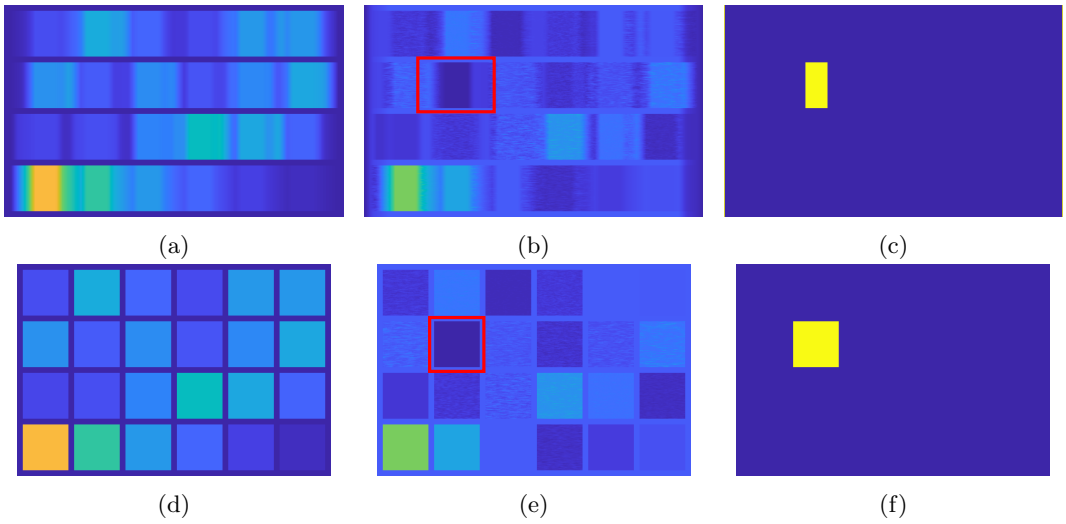


Figure 6.30: CASSI-TD predictions using a Boolean CA, where **a,d)** are the compressed image measurements, **b,e)** show the absolute difference between the image and target measurements and, **c,f)** show the resultant binary detections where $K = 10$ and $T = 0.99$ using SD-CASSI and DD-CASSI respectively.

Each of the examples displayed in this section so far, have shown the result of performing CASSI-TD on a single target from the colour checker synthetic dataset using an ensemble CA of a fixed size ($K = 10$). Aggregating the performance in detecting each target and varying the number of shots in an ensemble for each CA type, a set of metrics characterising the wider performance can be obtained, these are plotted in Figure 6.31.

Generally, as the number of shots, K, increases, the performance of the CASSI-TD approach improves. This can be attributed to the increased spacing between sampling points in each CA which results in fewer interactions between adjacent spectra. The results of performing

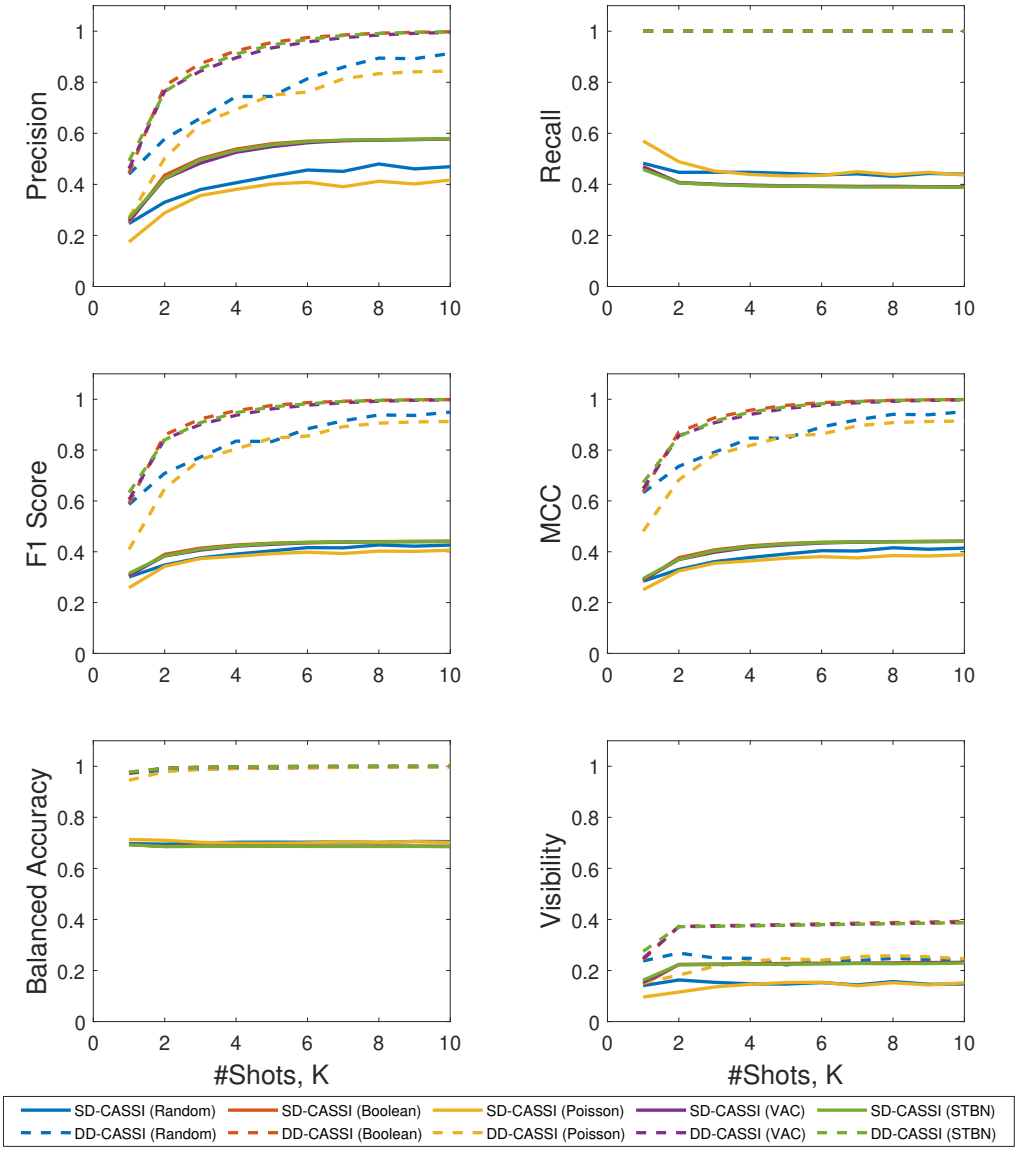


Figure 6.31: Per-shot performance metrics of CASSI-TD

CASSI-TD using an ensemble CA of increasing size (*i.e.,* an increasing number of shots, K) are displayed in Figure 6.32.

For $K = 1$, the transmittance of any CA is set as 0.5, meaning half of the elements allow light to be transmitted and half will block it completely. As a result, on average, only half of the target pixels can be identified in any target. From the precision curves in Figure 6.31 this can be observed for the DD-CASSI approach, where at $K = 1$, the precision achieved is approximately 0.5. The performance of the SD-CASSI approach is worse in every observed metric, due to the spatial shear obfuscating any target with components of the adjacent spectra. For $K \geq 2$, the transmittance is $1/K$ in each of the K shots. However, as highlighted in Figure 3.36, this results in every pixel being sampled when using a temporally constrained CA (*i.e.,* Boolean, VAC, and STBN), whereas the non-constrained CAs (*i.e.,* Random, and PD) can still have pixels in the image which are not sampled. This produces CASSI-TD results which are, on average, lower than using a constrained CA in both the SD-CASSI and DD-CASSI architectures.

As the erroneous detections are suppressed with additional shots, the "errors" which occur in the target region are removed which causes the dip in recall and balanced accuracy for each of the SD-CASSI and CAs combinations tested. This is also the cause of the slower rate of change in the F1 score and MCC compared to that of the DD-CASSI examples. The balanced accuracy for all CA types tested remains relatively constant with the use of either SD-CASSI or DD-CASSI architectures as, in both, the ability to differentiate the desired foreground from the remainder of the image is consistent over each shot.

The average target visibility is relatively poor compared to other TD approaches (Figures 5.20 and 5.27). This can be attributed to; a lack of negative spectra suppression, and materials with similar spectra, or with similar energy (*i.e.,* AUC as discussed above) appearing similar in the output of a CASSI system. This is shown in Figures 6.25a and 6.25b where materials with different spectra can appear similar in the resultant compressed measurement. Despite this, however, the performance of the proposed CASSI-TD approach is reasonable given the spectra of the desired target are never directly observed.

(a) (b) (c) (d) (e)

Figure 6.32: Example of CASSI-TD predictions with increasing CA ensemble size, where $K = [1, 2, 5, 7, 10]$.

6.4.1.3 Colour CASSI-TD Performance on Synthetic Macbeth Chart

The process described in Section 6.4.1.2 for testing the proposed TD procedure using the CASSI architectures was repeated for the simulated CCASSI architectures. The CCASSI approach has an additional design parameter in the set of, F, filters which replace the binary bandpass filters in each element of a binary CA. Two methods for generating these filters are considered and tested using the colour checker synthetic dataset.

A) Spectra-filling Filter Generation

The first method involved generating a set of non-intersecting filters which uniformly cover the range of wavelengths investigated. An example set of $F = 5$ filters along with the spectrum of target 8 of the colour checker dataset are shown in Figure 6.33, where each filter will sample the corresponding portion of the spectrum within its pass band.

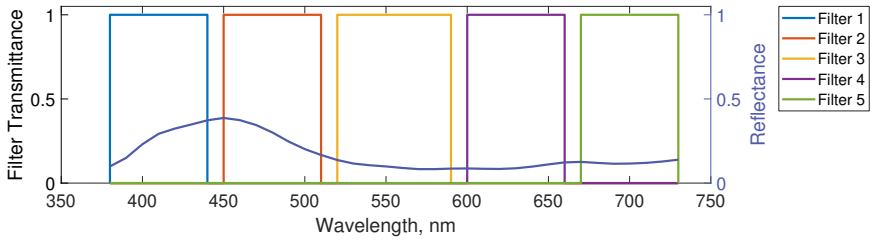


Figure 6.33: Spectral filters generated using uniform non-intersection rules.

B) WaLuMI Filter Generation

Alternatively, band selection techniques, in this case WaLuMI, can be used to identify the most informative wavelengths for each target spectra with the corresponding filters defined as some window of non-zero width, centred at each of these candidate wavelengths. The width of these filters can be defined with an additional parameter, $f_{padding}$, which controls the transmittance value of a colour CA. An example of this is shown in Figure 6.34 for target 8 of the colour checker dataset where $F = 5$ and $f_{padding} = 3$.

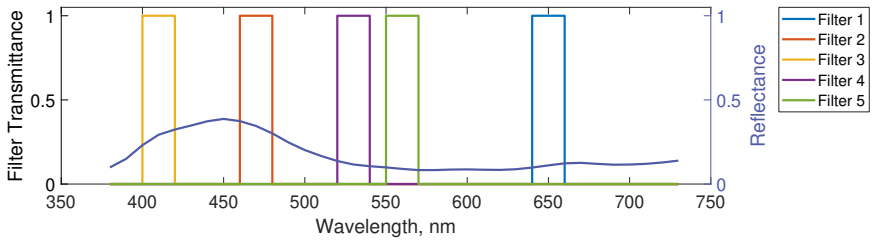


Figure 6.34: Candidate spectral filters generated by WaLuMI based on the spectral response of target 8 from Figure 6.24.

In order to compare the performance of colour CASSI-TD when using either the uniform or WaLuMI-derived colour filters, candidate bands were generated using WaLuMI with $F = 5$ and $f_{padding} = 3$ for each of the twenty-four target spectra in Figure 6.24. The average F1 score and MCC over each of the targets, at $T = 0.99$ and $K = 10$, using both single and dual-dispersion CASSI and CCASSI and both Random and Boolean CAs are shown in Table 6.2. Where the results of using both uniform and WaLuMI filters in the CCASSI architectures are displayed.

Table 6.2: Comparison between uniform and WaLuMI colour filter generation.

	SD-CASSI		DD-CASSI		SD-CCASSI				DD-CCASSI			
	Random	Boolean	Random	Boolean	Random		Boolean		Random		Boolean	
					Uniform	WaLuMI	Uniform	WaLuMI	Uniform	WaLuMI	Uniform	WaLuMI
F1	0.4423	0.4471	0.9991	1.0000	0.4418	0.4280	0.4424	0.4456	0.9990	0.9761	0.9977	0.9901
MCC	0.4438	0.4717	0.0991	1.0000	0.4402	0.4081	0.4529	0.4342	0.9990	0.9766	0.9976	0.9901

From Table 6.2 it is apparent that, similar to the CASSI architectures, the DD-CCASSI models outperform the equivalent single dispersion model considerably in a TD task. The use of WaLuMI filters in colour CASSI-TD models does not improve performance when compared to the results gained with uniform filters, however, it does significantly increase the computational cost of CA design. As a result, the non-overlapping, uniform, filters are used for the remainder of this chapter unless otherwise specified. Additional research into alternative methods for generating optimal colour CAs may provide improved results in other, non-synthetic, datasets.

The results of using colour CASSI-TD to detect the "purplish blue" target from Figure 6.23 are shown in Figure 6.35 using a Random CA, and Figure 6.36 using a Boolean CA. The result of using both SD-CCASSI and DD-CCASSI architectures are shown in each case.

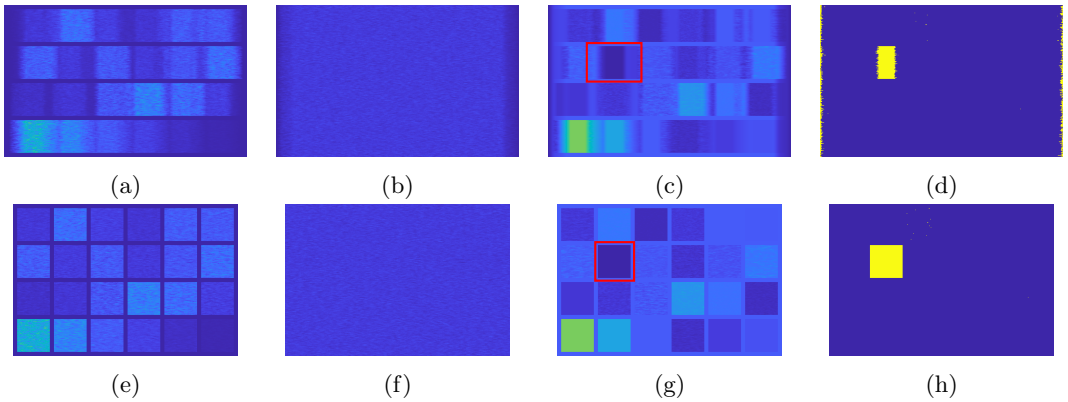


Figure 6.35: Colour CASSI-TD predictions using a Random CA, where **a,e)** are the compressed image measurements, **b,f)** are the compressed target image measurements, **c,g)** show the absolute difference between the image and target measurements, and **d,h)** show the resultant binary detections where $K = 10$ and $T = 0.99$ using SD-CCASSI and DD-CCASSI respectively.

Similar to the results presented in Section 6.4.1.2 for binary CASSI-TD, a DD-CCASSI architecture provides the best detection. The SD-CCASSI model achieves an F1 score of 0.4325

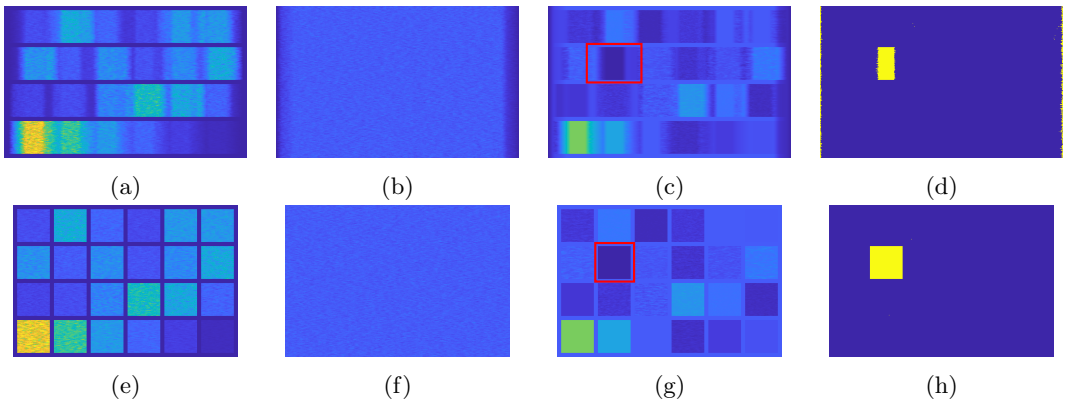


Figure 6.36: Colour CASSI-TD predictions using a Boolean CA, where **a,e)** are the compressed image measurements, **b,f)** are the compressed target image measurements, **c,g)** show the absolute difference between the image and target measurements, and **d,h)** show the resultant binary detections where $K = 10$ and $T = 0.99$ using SD-CCASSI and DD-CCASSI respectively.

and 0.4365 for Random and Boolean CAs respectively, whereas the DD-CCASSI achieves scores of 0.9983 and 0.9998. In each case, this is a slight reduction compared to those achieved in the binary CASSI-TD simulations.

Figure 6.37 shows the results of aggregating the performance metrics from detecting each target using each of the CA designs with both single and dual dispersion CCASSI. As the number of shots, K, is increased, the precision, F1 score, MCC, and visibility trend upwards, similar to the behaviour of the CASSI architectures. The recall and balanced accuracy of the SD-CCASSI models both decrease slightly for the same reasons as discussed for the SD-CASSI model. Both binary and colour CASSI-TD architectures are compared in Section 6.4.1.4.

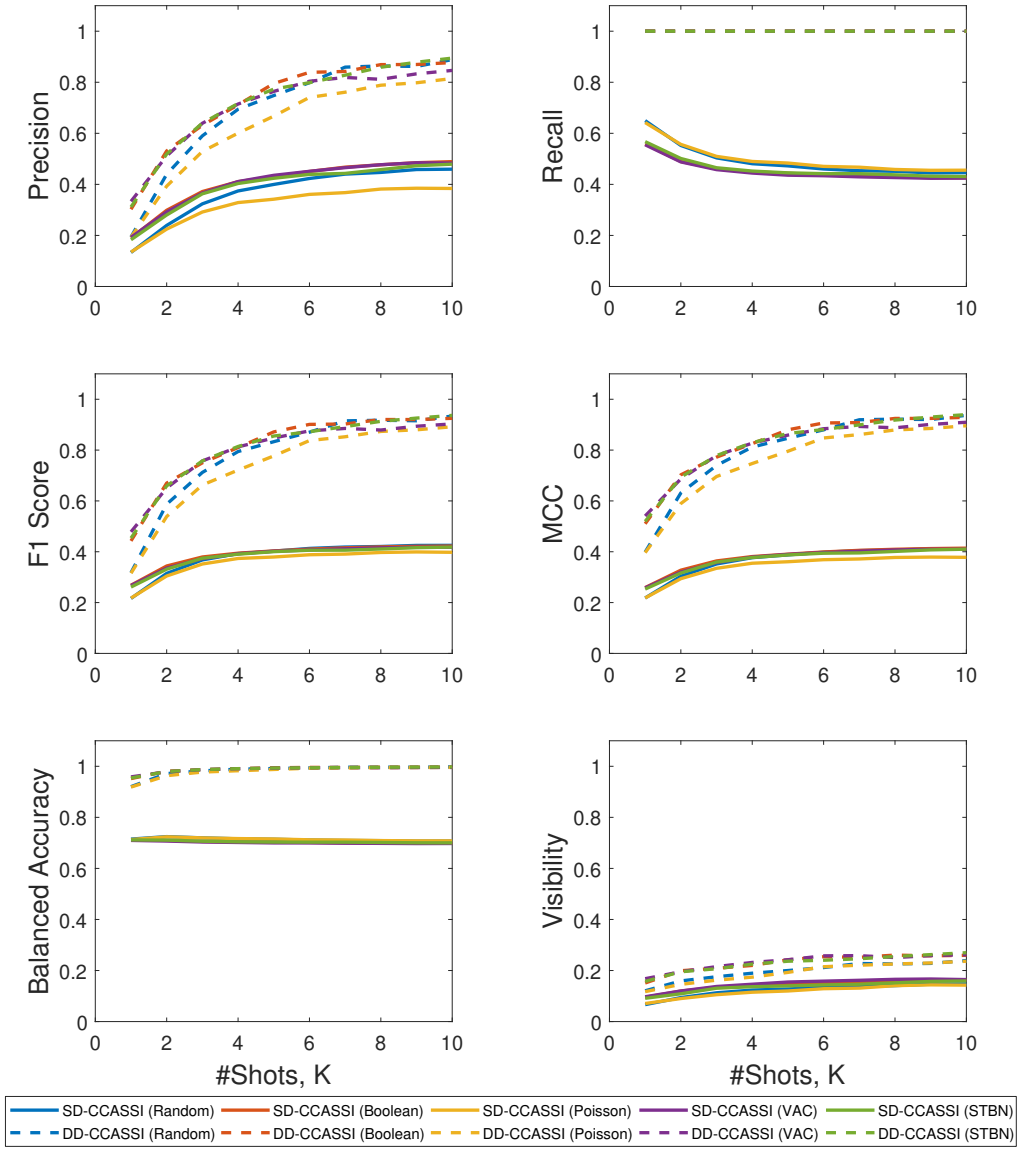


Figure 6.37: Per-shot performance metrics of colour CASSI-TD

6.4.1.4 Comparing the use of CASSI and CCASSI for TD

In the previous sections, the performance of both binary CASSI-TD and colour CASSI-TD have been discussed individually when applied to the synthetic colour checker image. In both cases, the dual dispersion architectures were shown to be superior, when considering TD.

From Figures 6.31 and 6.37 there is little difference between the structured and unstructured CA designs with Boolean, VAC, and STBN designs each performing similarly in both the colour and binary cases, whilst exceeding the performance of both Random and PD CAs. As a result, for the remainder of this chapter, the results of using Boolean CAs are shown. Boolean CAs are shown as, since there is little difference between the structured CAs, this will be representative of the results obtainable using VAC and STBN CAs. Additionally, it is the most accessible when considering the computation required to generate such CAs, as discussed in Section 3.7.4.

A comparison between the binary and colour CASSI-TD approaches with an increasing number of ensemble CA shots, K, is shown in Figure 6.38 using a grouped box plot. The box plots are made by collecting the F1 score, MCC, balanced accuracy, and visibility for each of the targets in the image, with the box representing the range between the 25th and 75th percentiles, and the circles represent outliers. The results for each architecture are presented for each of the shots, K, such that a comparison can be made in addition to the general trend as K increases.

As seen in Figures 6.31 and 6.37 for the individual binary or colour architectures, there is a general trend of improvement with an increasing number of shots. Comparing the performance of a binary CASSI architecture to the equivalent colour CASSI architecture it is noticeable that the CASSI architecture has a far better performance at lower values of K, notably $K = 1$. This is due to the lower transmittance of colour CAs at low values of K, and as a result there are potentially informative elements in the image not encoded in the compressed representation which will hamper the ability of the proposed CASSI-TD approach to properly isolate targets.

While the DD-CCASSI architecture can achieve similar F1 scores and MCC values as the DD-CASSI using larger CA ensembles, the latter achieves visibility far in excess of the former, indicating a strong ability to separate the target of interest from background material. Both dual dispersion architectures outperform the equivalent single-dispersion architectures by a large margin on average.

Finally, the results of applying the CASSI-TD approach for each target are tabulated for each architecture individually in Tables 6.3 to 6.6 for $K = [1, 5, 10]$. This aids in informing and expanding the trends observed in Figure 6.38, with both dual dispersion architectures achieving near-perfect detection at $K = 5$ and higher.

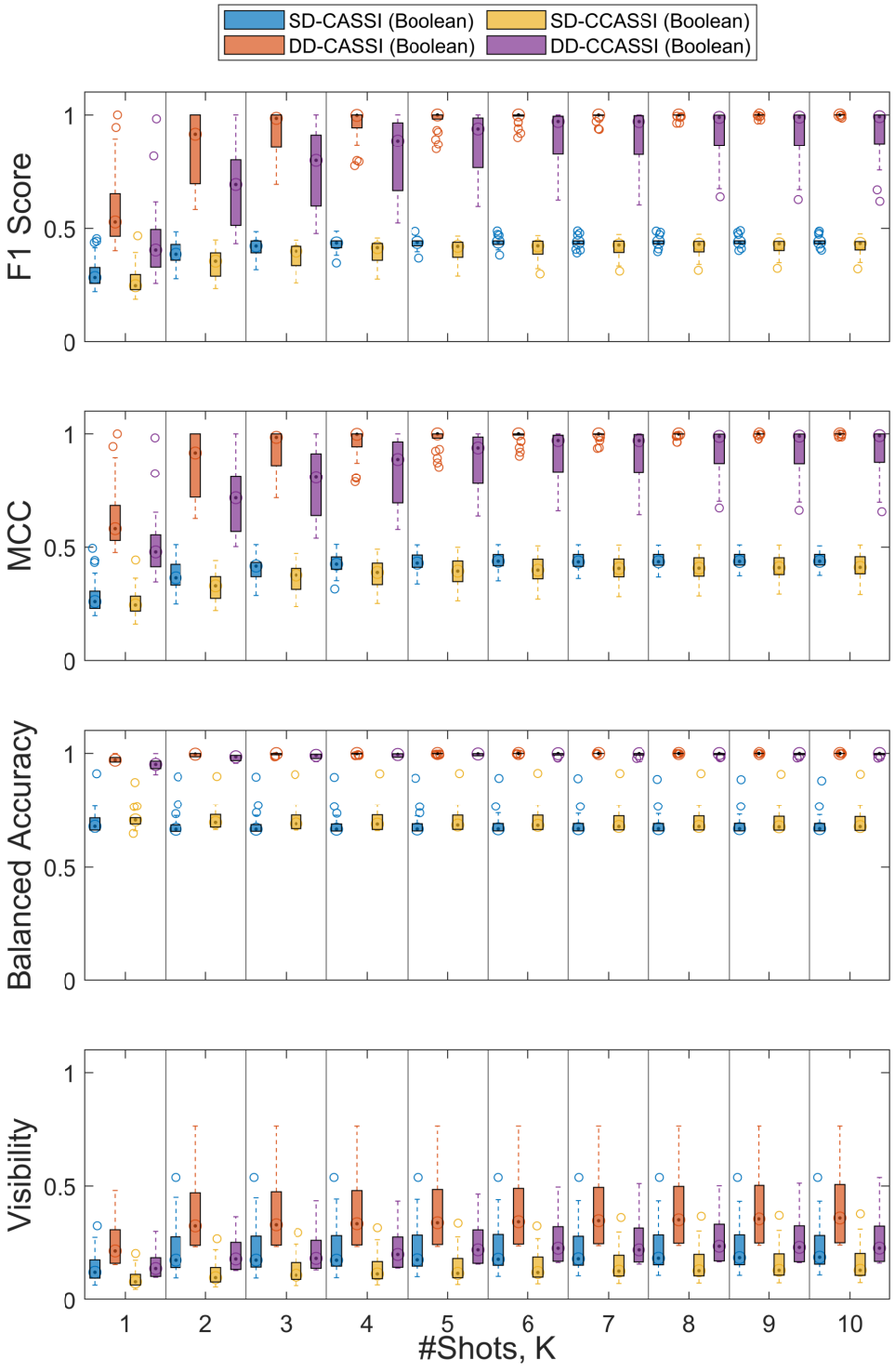


Figure 6.38: Per-shot comparison between binary and colour CASSI-TD approaches using a Boolean CA with a detection threshold of $T = 0.99$.

Table 6.3: Single dispersion binary CASSI-TD performance on the synthetic colour checker dataset.

SD-CASSI (Boolean)															
$T = 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.69	0.24	0.45	0.31	0.28	0.66	0.63	0.34	0.44	0.44	0.66	0.64	0.33	0.44	0.44
2	0.65	0.44	0.32	0.37	0.35	0.65	0.87	0.3	0.45	0.5	0.65	0.87	0.3	0.45	0.5
3	0.65	0.24	0.35	0.29	0.25	0.66	0.69	0.32	0.44	0.46	0.66	0.7	0.32	0.44	0.46
4	0.67	0.21	0.41	0.27	0.24	0.67	0.59	0.34	0.43	0.43	0.67	0.59	0.34	0.44	0.43
5	0.65	0.33	0.34	0.34	0.3	0.65	0.86	0.3	0.45	0.5	0.65	0.87	0.3	0.45	0.5
6	0.65	0.34	0.34	0.34	0.31	0.65	0.87	0.3	0.45	0.5	0.65	0.87	0.3	0.45	0.5
7	0.69	0.32	0.41	0.36	0.33	0.66	0.7	0.32	0.44	0.46	0.66	0.7	0.32	0.44	0.46
8	0.66	0.25	0.36	0.3	0.26	0.65	0.77	0.31	0.44	0.48	0.65	0.77	0.31	0.45	0.48
9	0.66	0.3	0.35	0.32	0.29	0.65	0.87	0.3	0.45	0.5	0.65	0.87	0.3	0.45	0.5
10	0.65	0.26	0.36	0.3	0.26	0.65	0.77	0.31	0.44	0.48	0.65	0.77	0.31	0.45	0.48
11	0.66	0.29	0.36	0.32	0.29	0.66	0.7	0.32	0.44	0.46	0.66	0.7	0.32	0.44	0.46
12	0.66	0.39	0.34	0.36	0.34	0.65	0.77	0.31	0.45	0.48	0.65	0.77	0.31	0.45	0.48
13	0.73	0.25	0.53	0.34	0.32	0.67	0.51	0.35	0.42	0.4	0.67	0.55	0.35	0.43	0.42
14	0.67	0.21	0.41	0.28	0.24	0.67	0.56	0.35	0.43	0.42	0.67	0.56	0.35	0.43	0.42
15	0.7	0.28	0.46	0.35	0.32	0.66	0.69	0.32	0.44	0.46	0.66	0.64	0.33	0.44	0.44
16	0.69	0.49	0.39	0.44	0.42	0.66	0.7	0.32	0.44	0.46	0.66	0.7	0.32	0.44	0.46
17	0.67	0.44	0.36	0.4	0.37	0.65	0.87	0.3	0.45	0.5	0.65	0.87	0.3	0.45	0.5
18	0.66	0.23	0.38	0.29	0.25	0.67	0.59	0.34	0.43	0.43	0.67	0.6	0.34	0.44	0.43
19	0.66	0.69	0.32	0.44	0.46	0.65	0.87	0.3	0.45	0.5	0.65	0.87	0.3	0.45	0.5
20	0.67	0.63	0.35	0.45	0.45	0.65	0.87	0.3	0.45	0.5	0.65	0.87	0.3	0.45	0.5
21	0.65	0.34	0.34	0.34	0.31	0.65	0.87	0.3	0.45	0.5	0.65	0.87	0.3	0.45	0.5
22	0.67	0.3	0.37	0.33	0.3	0.65	0.76	0.31	0.44	0.47	0.65	0.77	0.31	0.45	0.48
23	0.7	0.24	0.47	0.32	0.29	0.67	0.54	0.36	0.43	0.42	0.67	0.56	0.35	0.43	0.42
24	0.88	0.35	0.84	0.49	0.51	0.87	0.39	0.79	0.52	0.53	0.85	0.43	0.75	0.55	0.54
Average	0.68	0.34	0.4	0.35	0.32	0.67	0.72	0.34	0.44	0.47	0.67	0.73	0.34	0.45	0.47

BAcc - Balanced Accuracy, P - Precision, R - Recall, F1 - F1 Score, MCC - Matthews Correlation Coefficient.

Table 6.4: Single dispersion colour CASSI-TD performance on the synthetic colour checker dataset.

SD-CCASSI (Boolean)															
$T = 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.73	0.19	0.58	0.29	0.28	0.7	0.45	0.43	0.44	0.42	0.68	0.55	0.38	0.45	0.44
2	0.68	0.31	0.4	0.35	0.32	0.66	0.76	0.32	0.45	0.48	0.65	0.8	0.31	0.45	0.49
3	0.65	0.19	0.39	0.25	0.22	0.67	0.55	0.35	0.43	0.42	0.66	0.63	0.34	0.44	0.44
4	0.7	0.17	0.52	0.25	0.24	0.68	0.41	0.39	0.4	0.37	0.68	0.5	0.37	0.42	0.41
5	0.67	0.27	0.39	0.32	0.28	0.66	0.76	0.32	0.45	0.48	0.65	0.83	0.31	0.45	0.49
6	0.7	0.32	0.44	0.37	0.34	0.66	0.69	0.32	0.44	0.46	0.65	0.78	0.31	0.44	0.48
7	0.72	0.24	0.53	0.33	0.31	0.68	0.55	0.37	0.44	0.43	0.67	0.61	0.34	0.44	0.44
8	0.66	0.18	0.41	0.25	0.22	0.67	0.58	0.34	0.43	0.43	0.66	0.63	0.33	0.44	0.44
9	0.68	0.22	0.44	0.29	0.27	0.66	0.69	0.32	0.44	0.46	0.65	0.76	0.31	0.44	0.47
10	0.67	0.17	0.43	0.24	0.22	0.66	0.63	0.33	0.43	0.43	0.66	0.71	0.32	0.44	0.46
11	0.69	0.24	0.46	0.31	0.28	0.67	0.59	0.35	0.44	0.43	0.66	0.62	0.34	0.44	0.44
12	0.7	0.28	0.46	0.35	0.32	0.66	0.6	0.34	0.43	0.43	0.66	0.66	0.33	0.44	0.45
13	0.75	0.2	0.62	0.31	0.3	0.73	0.42	0.5	0.45	0.43	0.72	0.44	0.46	0.45	0.42
14	0.69	0.18	0.5	0.26	0.24	0.69	0.43	0.41	0.42	0.39	0.68	0.48	0.39	0.43	0.41
15	0.72	0.2	0.54	0.29	0.28	0.7	0.46	0.41	0.44	0.41	0.68	0.53	0.37	0.44	0.42
16	0.7	0.32	0.45	0.37	0.34	0.67	0.55	0.35	0.43	0.42	0.67	0.6	0.34	0.44	0.44
17	0.69	0.3	0.44	0.36	0.33	0.66	0.76	0.32	0.45	0.48	0.65	0.83	0.31	0.45	0.49
18	0.71	0.21	0.51	0.3	0.28	0.68	0.52	0.37	0.43	0.42	0.67	0.56	0.35	0.43	0.42
19	0.71	0.52	0.43	0.47	0.45	0.66	0.74	0.32	0.45	0.47	0.65	0.79	0.31	0.45	0.48
20	0.69	0.44	0.4	0.42	0.4	0.66	0.76	0.32	0.45	0.48	0.65	0.8	0.31	0.45	0.48
21	0.67	0.27	0.4	0.32	0.29	0.65	0.78	0.31	0.44	0.48	0.65	0.81	0.31	0.44	0.48
22	0.67	0.21	0.42	0.28	0.25	0.66	0.64	0.33	0.43	0.44	0.66	0.68	0.33	0.44	0.45
23	0.72	0.18	0.55	0.27	0.26	0.69	0.41	0.4	0.4	0.37	0.68	0.48	0.37	0.42	0.4
24	0.88	0.27	0.87	0.41	0.45	0.88	0.35	0.83	0.49	0.51	0.87	0.36	0.81	0.5	0.51
Average	0.7	0.25	0.48	0.32	0.3	0.68	0.59	0.38	0.44	0.44	0.67	0.64	0.36	0.44	0.45

BAcc - Balanced Accuracy, P - Precision, R - Recall, F1 - F1 Score, MCC - Matthews Correlation Coefficient.

Table 6.5: Dual dispersion binary CASSI-TD performance on the synthetic colour checker dataset.

DD-CASSI (Boolean)															
$T = 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1.00	0.98	0.41	1.00	0.58	0.62	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	0.99	0.68	1.00	0.81	0.82	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3	0.98	0.43	1.00	0.6	0.64	1.00	0.98	1.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00
4	0.98	0.42	1.00	0.59	0.63	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
5	0.98	0.51	1.00	0.68	0.7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
6	0.98	0.5	1.00	0.67	0.69	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	0.99	0.55	1.00	0.71	0.73	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
8	0.98	0.52	1.00	0.68	0.71	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
9	0.99	0.55	1.00	0.71	0.73	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
10	0.98	0.48	1.00	0.65	0.68	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
11	0.99	0.61	1.00	0.76	0.77	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
12	0.99	0.66	1.00	0.79	0.81	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
13	0.98	0.47	1.00	0.64	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
14	0.98	0.44	1.00	0.61	0.65	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
15	0.99	0.61	1.00	0.76	0.77	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
16	1.00	0.9	1.00	0.95	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
17	0.99	0.65	1.00	0.79	0.8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
18	0.98	0.41	1.00	0.58	0.63	1.00	0.99	1.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00
19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
20	1.00	0.94	1.00	0.97	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
21	0.98	0.49	1.00	0.66	0.69	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
22	0.98	0.48	1.00	0.65	0.68	1.00	0.99	1.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00
23	0.99	0.66	1.00	0.79	0.8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
24	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Average	0.99	0.6	1.00	0.73	0.76	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

BAcc - Balanced Accuracy, P - Precision, R - Recall, F1 - F1 Score, MCC - Matthews Correlation Coefficient.

Table 6.6: Dual dispersion colour CASSI-TD performance on the synthetic colour checker dataset.

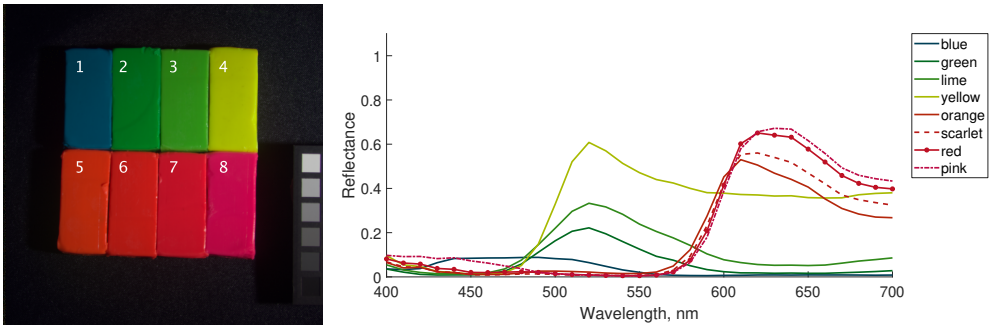
DD-CCASSI (Boolean)															
$T = 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.96	0.29	1.00	0.45	0.52	1.00	0.86	1.00	0.92	0.92	1.00	0.96	1.00	0.98	0.98
2	0.98	0.52	1.00	0.68	0.71	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3	0.97	0.33	1.00	0.5	0.56	1.00	0.93	1.00	0.96	0.96	1.00	1.00	1.00	1.00	1.00
4	0.96	0.29	1.00	0.45	0.52	1.00	0.87	1.00	0.93	0.93	1.00	0.96	1.00	0.98	0.98
5	0.98	0.46	1.00	0.63	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
6	0.98	0.49	1.00	0.66	0.69	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	0.98	0.42	1.00	0.59	0.63	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
8	0.97	0.32	1.00	0.49	0.55	1.00	0.97	1.00	0.99	0.98	1.00	1.00	1.00	1.00	1.00
9	0.98	0.4	1.00	0.57	0.62	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
10	0.96	0.31	1.00	0.47	0.53	1.00	0.95	1.00	0.97	0.97	1.00	0.99	1.00	1.00	1.00
11	0.98	0.44	1.00	0.61	0.65	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
12	0.98	0.5	1.00	0.66	0.69	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
13	0.97	0.37	1.00	0.54	0.59	1.00	0.97	1.00	0.99	0.98	1.00	1.00	1.00	1.00	1.00
14	0.97	0.32	1.00	0.49	0.55	1.00	0.98	1.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00
15	0.98	0.4	1.00	0.57	0.62	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
16	0.99	0.62	1.00	0.76	0.78	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
17	0.98	0.5	1.00	0.67	0.7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
18	0.97	0.33	1.00	0.5	0.56	1.00	0.95	1.00	0.98	0.98	1.00	1.00	1.00	1.00	1.00
19	1.00	0.98	1.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
20	1.00	0.82	1.00	0.9	0.9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
21	0.98	0.44	1.00	0.61	0.65	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
22	0.97	0.37	1.00	0.54	0.59	1.00	0.96	1.00	0.98	0.98	1.00	1.00	1.00	1.00	1.00
23	0.97	0.38	1.00	0.55	0.6	1.00	0.96	1.00	0.98	0.98	1.00	1.00	1.00	1.00	1.00
24	0.98	0.51	1.00	0.67	0.7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Average	0.98	0.45	1.00	0.61	0.65	1.00	0.97	1.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00

BAcc - Balanced Accuracy, P - Precision, R - Recall, F1 - F1 Score, MCC - Matthews Correlation Coefficient.

6.4.2 CASSI-TD Results on the CAVE MSI dataset

While the synthetic image of a colour checker provides a good way of benchmarking the behaviours of each of the modelled CASSI architectures when applied to a TD task, it is highly idealised and does not exhibit the natural variation present in non-synthetic imagery. In order to test the performance of the proposed CASSI-TD approach on more natural images, a subset of the CAVE multispectral image dataset [291] which includes objects suitable for TD, *i.e.,* those with a series of small to medium objects with fairly uniform spectra, were selected. The images included in this subset are shown in Appendix A.2.

The example image used in this section is shown in Figure 6.39a, where the target objects are 8 slabs of coloured clay, the spectra of which (obtained by averaging a small region of each slab) are shown in Figure 6.39b. These targets are numbered from left to right and top to bottom for ease of reference. The images from the CAVE MSI dataset are all $M = N = 512$ and $L = 31$, in the range $\lambda = 400 - 1000$ nm.



(a) (b)

Figure 6.39: **a)** Example image from the CAVE MSI dataset with overlaid target numbers. **b)** Target spectra present in Figure 6.39a.

The first target from Figure 6.39a used to highlight the result of performing CASSI-TD is target 4, "yellow clay". From the compressed measurements obtained using each of the simulated CASSI architectures, shown in Figures 6.40a to 6.40d, it is apparent that this is indeed one of, if not, the most easily separable targets in the image. This can be attributed to its easily distinguishable spectrum and spectral density (AUC) from Figure 6.39b.

The absolute differences between the compressed measurement of the scene, Figures 6.40a to 6.40d, and the compressed measurement based on the generated target image for each of the four CASSI architectures are shown in Figures 6.40e to 6.40h respectively. From these images, there is a good degree of separation, with the DD-CASSI model achieving a visibility of 0.71 compared to 0.48, 0.31, and 0.41 for the SD-CASSI, SD-CCASSI, and DD-CCASSI

architectures respectively. This results in similarly accurate detections in the binary images shown in Figures 6.40i to 6.40l, where applying a threshold of $T = 0.9$, F1 scores of 0.4, 0.48, 0.71, and 0.84 are achieved for the SD-CASSI, SD-CCASSI, DD-CASSI, and DD-CCASSI models respectively. Despite retaining similar proportions of the target compared to the dual dispersion architectures, the single dispersion architectures have their performance reduced by the low energy strips at the edges of the compressed measurement resulting in lower precision. The CCASSI architectures achieve slightly higher F1 scores in each case at this threshold than the equivalent binary CASSI, however, on average they achieve poorer results across each of the targets considered over a range of thresholds (as displayed in Table 6.7) indicating they are much more sensitive to the choice of threshold, T.

The second target from Figure 6.39a selected to investigate the performance of CASSI-TD is target 6, "scarlet clay". From Figures 6.40a to 6.40d, this material appears much more similar to others in the compressed measurement when compared to the previous target considered. With visibilities of 0.32, 0.28, 0.36 and 0.27 for the SD-CASSI, SD-CCASSI, DD-CASSI, and DD-CCASSI models, respectively, a notable decrease compared to those achieved with target 4. The absolute difference between the compressed measurements obtained using each CASSI architecture and the synthesised target image are shown in Figures 6.40m to 6.40p and the thresholded images, with $T = 0.95$, for target 6 are shown in Figures 6.40q to 6.40t achieving F1 scores of 0.43, 0.45, 0.57, and 0.51 respectively.

Similar to the results obtained by performing CASSI-TD on the spectra for target 4, the dual dispersion methods show improved performance compared with the single dispersion architectures. In addition, the CCASSI architectures retain a larger proportion of the target material, with SD-CCASSI offering a slight improvement over the equivalent SD-CASSI model when considering F1 score. However, each of the architectures, most notably the DD-CCASSI, have a high FPR including a wider range of similar materials in the predicted target map, highlighting its sensitivity to colour and illumination variations.

The effect of changing the size of the ensemble Boolean CA between $K = 1$ and $K = 10$ on the average performance of each of the simulated CASSI architectures is shown in the grouped box plot in Figure 6.41. The trend of improving with an increasing number of shots is once again observed, with the DD-CASSI architecture achieving, on average, the highest values in each of the metrics used.

The results for each individual target using a Boolean ensemble CA with $K = [1, 5, 10]$ averaged using thresholds between $T = [0.9 - 0.99]$ are shown in Table 6.7. The DD-CASSI architecture achieves the highest performance of each of the simulated CASSI models, however,

the DD-CCASSI consistently achieves a higher recall value, indicating an increased resilience to slight colour changes in this image at the expense of a loss in precision. Similarly, the SD-CCASSI achieves higher recall than the equivalent SD-CASSI model. Both single dispersion architectures suffer from relatively poor performance when compared to the dual dispersion architectures due to the shearing operation.

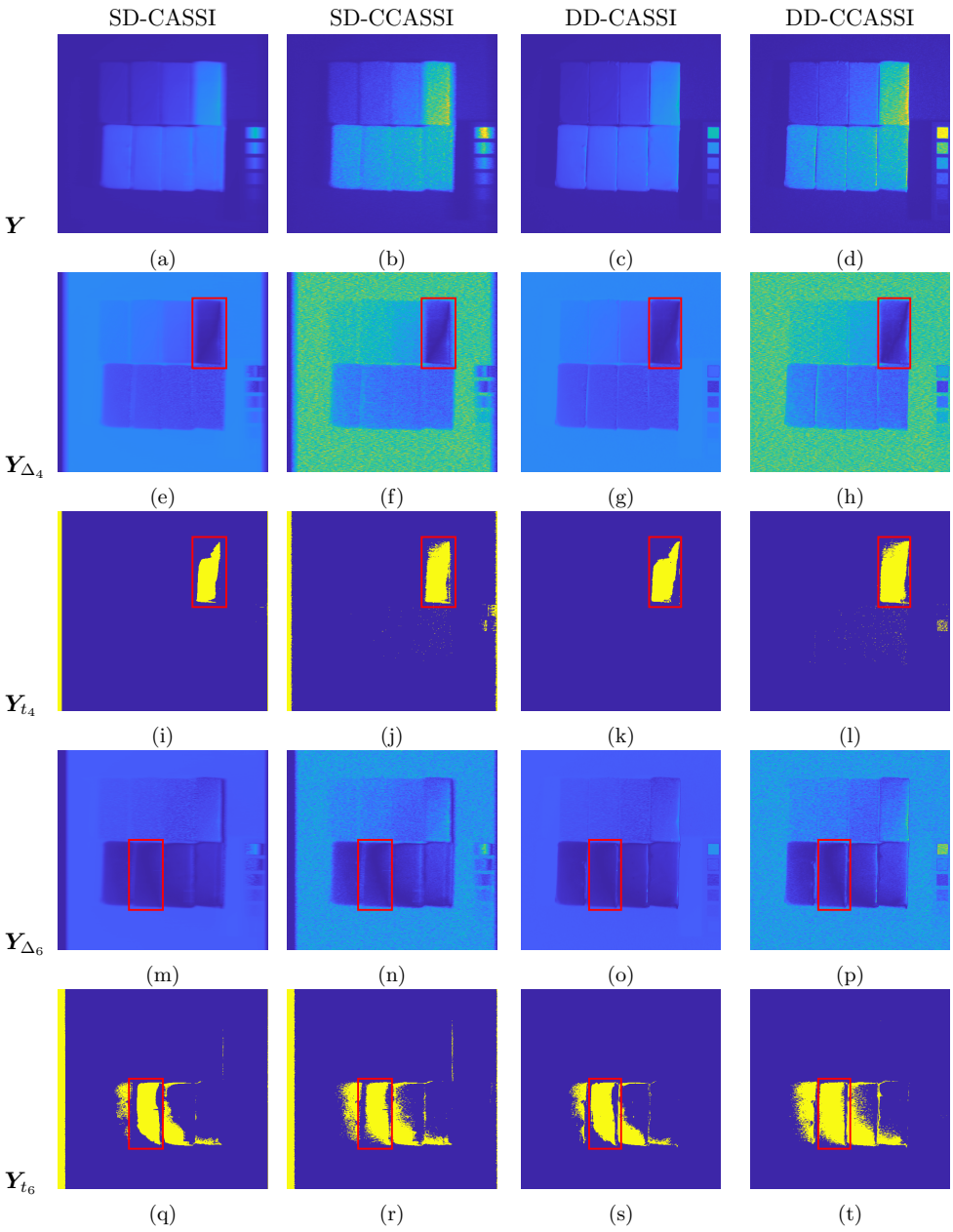


Figure 6.40: Example CASSI-TD results on the CAVE MSI dataset, each column shows the result using SD-CASSI, SD-CCASSI, DD-CASSI, and DD-CCASSI respectively. **a-d)** Compressed measurement of the scene from Figure 6.39a. **e-h)** Absolute difference between image measurement and target 4 image measurement. **i-l)** CASSI-TD predictions for target 4 with $K = 10$ and $T = 0.9$. **m-p)** Absolute difference between image measurement and target 6 image measurement. **q-t)** CASSI-TD predictions for target 6 with $K = 10$ and $T = 0.95$.

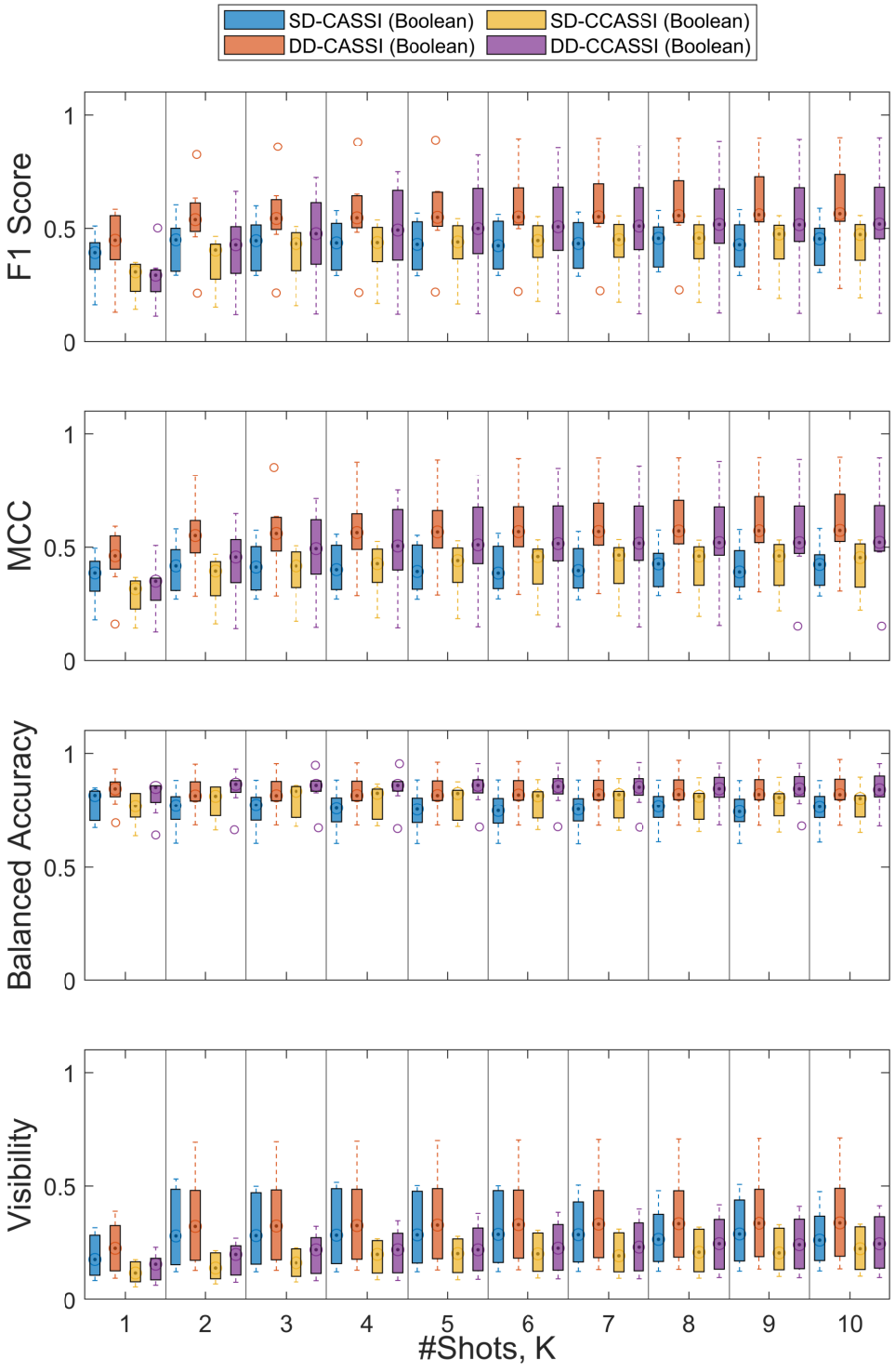


Figure 6.41: Per-shot comparison between binary and colour CASSI-TD approaches using a Boolean CA with a detection threshold of $T = 0.95$.

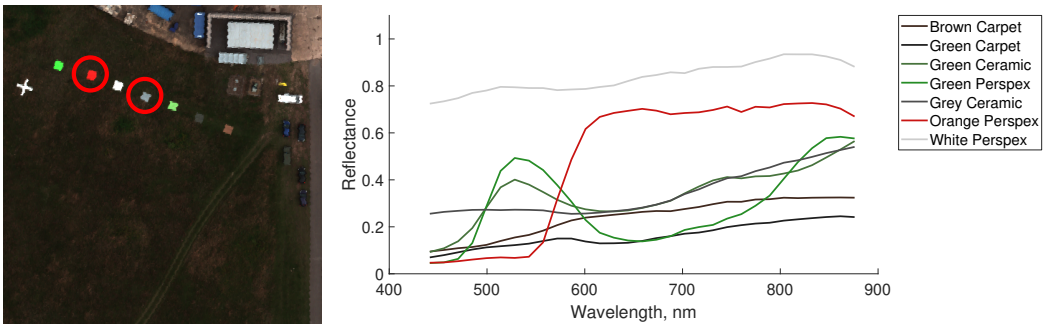
Table 6.7: CASSI-TD performance on the "clay" image from the CAVE MSI dataset.

SD-CASSI (Boolean)															
$T = 0.9 - 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.73	0.16	0.92	0.25	0.26	0.76	0.25	0.84	0.34	0.34	0.76	0.26	0.82	0.35	0.35
2	0.79	0.23	0.85	0.33	0.35	0.84	0.38	0.76	0.49	0.49	0.84	0.40	0.75	0.51	0.50
3	0.76	0.28	0.63	0.37	0.36	0.73	0.41	0.49	0.44	0.41	0.72	0.41	0.48	0.43	0.40
4	0.65	0.33	0.34	0.32	0.29	0.59	0.40	0.20	0.25	0.24	0.59	0.39	0.21	0.26	0.24
5	0.69	0.21	0.51	0.29	0.26	0.66	0.29	0.38	0.30	0.26	0.66	0.30	0.39	0.31	0.27
6	0.79	0.29	0.70	0.39	0.38	0.72	0.37	0.51	0.38	0.36	0.73	0.36	0.53	0.38	0.37
7	0.78	0.33	0.67	0.41	0.40	0.70	0.43	0.45	0.37	0.36	0.73	0.41	0.52	0.40	0.39
8	0.79	0.44	0.64	0.48	0.47	0.71	0.61	0.44	0.45	0.46	0.73	0.56	0.51	0.46	0.46
Average	0.75	0.28	0.66	0.36	0.35	0.71	0.39	0.51	0.38	0.37	0.72	0.39	0.53	0.39	0.37
SD-CCASSI (Boolean)															
$T = 0.9 - 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.68	0.11	0.94	0.18	0.19	0.73	0.18	0.90	0.27	0.27	0.75	0.21	0.87	0.30	0.31
2	0.73	0.14	0.90	0.23	0.24	0.81	0.27	0.83	0.38	0.39	0.83	0.32	0.81	0.43	0.44
3	0.75	0.19	0.73	0.29	0.29	0.77	0.39	0.60	0.45	0.44	0.76	0.41	0.56	0.46	0.44
4	0.69	0.24	0.49	0.31	0.27	0.65	0.41	0.34	0.36	0.33	0.63	0.42	0.29	0.33	0.31
5	0.69	0.15	0.62	0.24	0.21	0.69	0.26	0.47	0.31	0.28	0.68	0.28	0.44	0.31	0.28
6	0.78	0.19	0.80	0.30	0.31	0.77	0.32	0.64	0.40	0.39	0.76	0.34	0.60	0.40	0.39
7	0.79	0.21	0.78	0.32	0.33	0.78	0.37	0.64	0.42	0.42	0.75	0.39	0.57	0.41	0.40
8	0.79	0.24	0.77	0.35	0.36	0.79	0.48	0.64	0.49	0.49	0.77	0.54	0.58	0.49	0.49
Average	0.74	0.18	0.75	0.28	0.28	0.75	0.34	0.63	0.38	0.38	0.74	0.37	0.59	0.39	0.38
DD-CASSI (Boolean)															
$T = 0.9 - 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.76	0.17	0.97	0.26	0.29	0.81	0.35	0.94	0.40	0.43	0.82	0.41	0.94	0.45	0.47
2	0.84	0.30	0.92	0.39	0.42	0.91	0.61	0.88	0.65	0.67	0.92	0.67	0.88	0.69	0.72
3	0.86	0.42	0.79	0.51	0.53	0.85	0.91	0.71	0.76	0.77	0.85	0.95	0.70	0.78	0.79
4	0.74	0.59	0.51	0.51	0.51	0.66	1.00	0.32	0.46	0.53	0.66	1.00	0.32	0.46	0.53
5	0.80	0.30	0.68	0.39	0.40	0.76	0.49	0.56	0.45	0.46	0.76	0.57	0.55	0.48	0.49
6	0.82	0.32	0.72	0.41	0.42	0.78	0.50	0.61	0.47	0.48	0.78	0.55	0.61	0.49	0.51
7	0.80	0.36	0.66	0.43	0.44	0.75	0.58	0.54	0.49	0.50	0.76	0.63	0.54	0.50	0.52
8	0.80	0.47	0.65	0.50	0.50	0.76	0.75	0.52	0.56	0.58	0.76	0.78	0.53	0.57	0.59
Average	0.80	0.37	0.74	0.42	0.44	0.79	0.65	0.63	0.53	0.55	0.79	0.70	0.63	0.55	0.58
DD-CCASSI (Boolean)															
$T = 0.9 - 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.70	0.09	0.98	0.16	0.19	0.76	0.21	0.97	0.28	0.31	0.77	0.25	0.96	0.31	0.34
2	0.75	0.13	0.95	0.21	0.25	0.84	0.37	0.92	0.43	0.47	0.86	0.44	0.92	0.49	0.53
3	0.83	0.23	0.87	0.33	0.37	0.88	0.72	0.79	0.69	0.71	0.88	0.82	0.77	0.75	0.76
4	0.81	0.36	0.69	0.45	0.45	0.76	0.93	0.52	0.63	0.66	0.74	0.98	0.49	0.62	0.66
5	0.79	0.18	0.78	0.29	0.31	0.81	0.39	0.69	0.44	0.45	0.80	0.43	0.65	0.45	0.47
6	0.81	0.19	0.82	0.29	0.33	0.83	0.39	0.73	0.45	0.47	0.82	0.43	0.70	0.46	0.48
7	0.82	0.20	0.81	0.31	0.34	0.81	0.46	0.67	0.49	0.50	0.79	0.51	0.63	0.49	0.50
8	0.83	0.24	0.81	0.35	0.38	0.82	0.60	0.67	0.56	0.57	0.80	0.66	0.62	0.57	0.58
Average	0.79	0.20	0.84	0.30	0.33	0.81	0.51	0.75	0.50	0.52	0.81	0.57	0.72	0.52	0.54

BAcc - Balanced Accuracy, P - Precision, R - Recall, F1 - F1 Score, MCC - Matthews Correlation Coefficient.

6.4.3 CASSI-TD Results on the UDRC dataset

The final dataset used to test the performance of CASSI-TD was the subset of the UDRC imagery (Appendix A.3.2) used in Chapter 5, spectrally downsampled to have dimensions equal to the imagery in the CAVE data, *i.e.,* $512 \times 512 \times 31$. An example image from this dataset is shown in Figure 6.42a. The scene contains seven square tiles of differing colours and materials used as the targets of interest, with the spectrum of each tile plotted in Figure 6.42b.



(a) (b)

Figure 6.42: **a)** Example image from the cropped UDRC dataset. **b)** Target spectra present in Figure 6.42a.

The results of performing TD on the compressed measurements from each of the simulated CASSI architectures on two of the targets from Figure 6.42 are presented in this section. Specifically, target 6, "orange perspex", and target 5, "grey ceramic", circled in Figure 6.42a. The compressed measurements of the scene shown Figure 6.42a are displayed in Figures 6.43a to 6.43d for each of the simulated CASSI architectures.

The first target considered, "orange perspex", is very distinct from the background with visibilities of 0.18, 0.14, 0.55, and 0.41 achieved for the SD-CASSI, SD-CCASSI, DD-CASSI, and DD-CCASSI models respectively. This indicates that this target can more easily be identified by the dual dispersion architectures. Each of the targets are roughly 15 pixels wide and the number of bands in the image is $L = 31$. As a result, when using either the SD-CASSI or SD-CCASSI architectures, the energy from a single pixel is spread across the adjacent 30 pixels to the left or right depending on the orientation of the shearing operation. In this case, the shear is modelled as shifting the energy of pixels to the right of the image domain. The lower visibilities achieved using the single dispersion architectures can be attributed to the targets being smaller than the effective size of the smear produced by the single dispersive element, as a result, no areas of the target are free of interference from the adjacent background materials in the compressed measurement. This results in the single dispersion architectures being unable

to detect this target in the image, as shown in the binary detection images from Figures 6.43i and 6.43j. Despite this, by lowering the detection threshold, the desired target can be included in the set of detected pixels, however, with an increase in FPR. Conversely, the target was detected when using either of the dual dispersion architectures (Figures 6.43k and 6.43l), with F1 scores of 0.66 and 0.71 achieved with a threshold of $T = 0.95$. This disparity between the single and dual dispersion architectures is consistent in the remaining targets, as shown in Table 6.8 and Figure 6.44.

The second example target displayed in Figure 6.43, "grey ceramic", is less distinct with a visibility of 0.14, 0.12, 0.23, and 0.18 for the SD-CASSI, SD-CCASSI, DD-CASSI, and DD-CCASSI models respectively. Similarly to the binary detections from Figures 6.43i and 6.43j, the single dispersion architectures are unable to detect the target material in the scene. The reduction in target visibility is also reflected in the predictions shown in Figures 6.43s and 6.43t from DD-CASSI and DD-CCASSI respectively, where a number of FP detections are introduced. Due to the lower visibility, the detection threshold was raised to 0.97. Of the two dual dispersion architectures, the binary DD-CASSI architecture performs better, achieving an average F1 score of 0.4, whereas the DD-CCASSI architecture only achieves a score of 0.11. This reduction highlights the sensitivity of the CASSI-TD algorithm to slight changes in colour, with the FPs belonging to visually similar objects.

Comparisons between the four architectures, with an increasing number of shots using a Boolean CA, are shown in Figure 6.44, where $T = 0.95$. Additionally, the detection metrics for each target, and the average performance of all of the considered CASSI-TD architectures using a Boolean CA with $K = [1, 5, 10]$ are shown in Table 6.8. The values from Table 6.8 are obtained by averaging the detection metrics over the range of thresholds from $T = 0.9 - 0.99$ where the target numbers correspond to those given in Table 5.5.

From Table 6.8, targets 3 through 7 can be detected reliably to some degree, indicated by the high recall values. However, both target 1 and 2 ("brown carpet" and "green carpet") are only detected along with a high FPR, *i.e.,* high recall but low precision. From Figure 6.42a it is evident that not only are these targets the most similar to the background, but they are also the two targets with the lowest energy (AUC) and therefore will convey a limited amount of information in the compressed measurement using a CASSI architecture. As the number of shots increases, the performance measures trend upwards as shown in the grouped box plot in Figure 6.44, this is consistent with the behaviour exhibited when performing CASSI-TD on the other datasets (Figures 6.38 and 6.41). Once again, on average, the DD-CASSI architecture performs the best in all metrics considered. However, there are instances where DD-CCASSI

performs similarly or slightly better when considering the target recall, but introduces an increased number of FP detections, as reflected in the decreased precision, MCC, and F1 score. Of the three datasets tested, CASSI-TD achieves the lowest overall metrics on the UDRC data, this is due to the relative difficulty of the task and may benefit from improved detection algorithms.

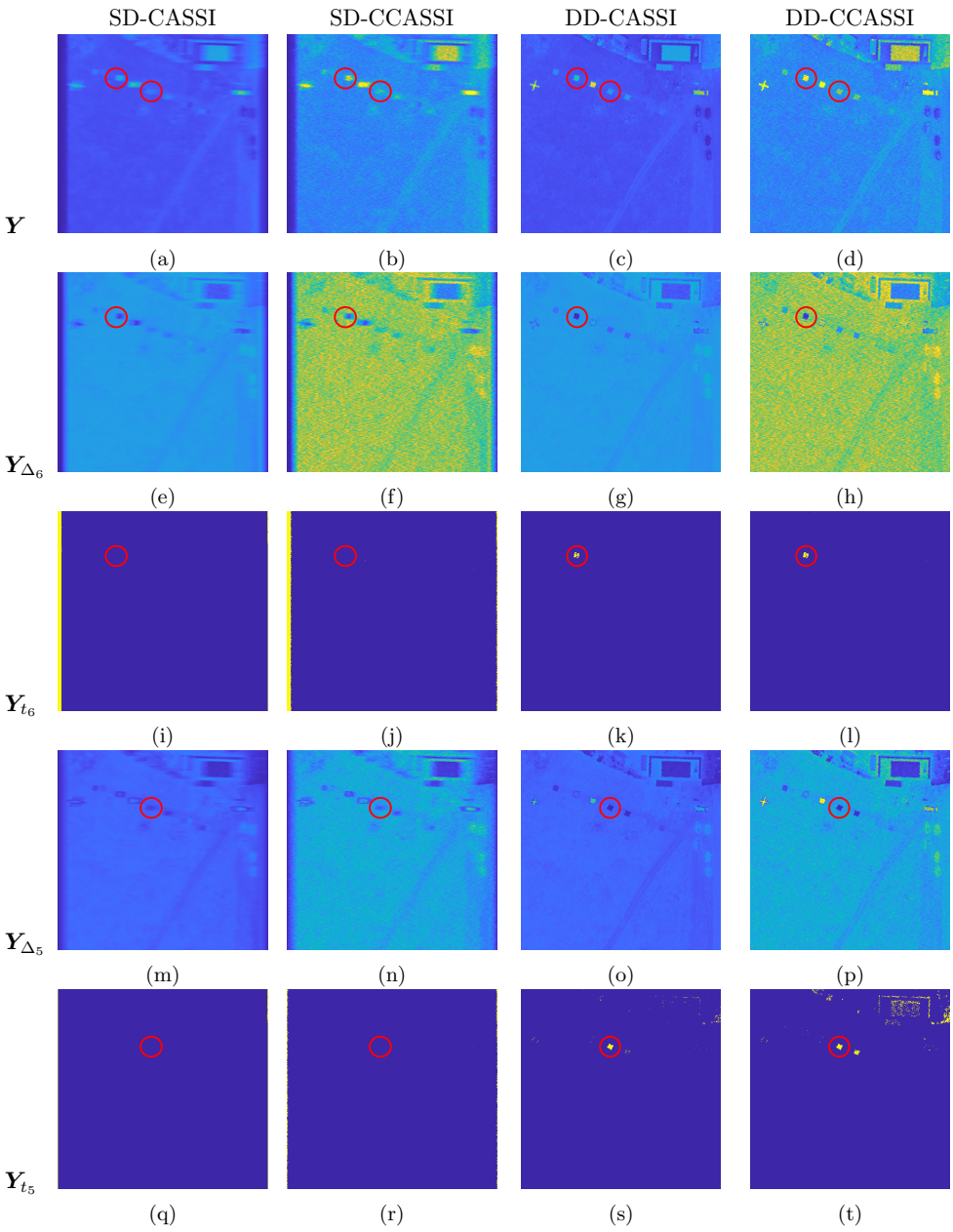


Figure 6.43: Example CASSI-TD results on the UDRC dataset, each column shows the result using SD-CASSI, SD-CCASSI, DD-CASSI, and DD-CCASSI respectively. **a-d)** Compressed measurement of the scene from Figure 6.42a. **e-h)** Absolute difference between image measurement and target 6 image measurement. **i-l)** CASSI-TD predictions for target 6 with $K = 10$ and $T = 0.95$. **m-p)** Absolute difference between image measurement and target 5 image measurement. **q-t)** CASSI-TD predictions for target 5 with $K = 10$ and $T = 0.97$.

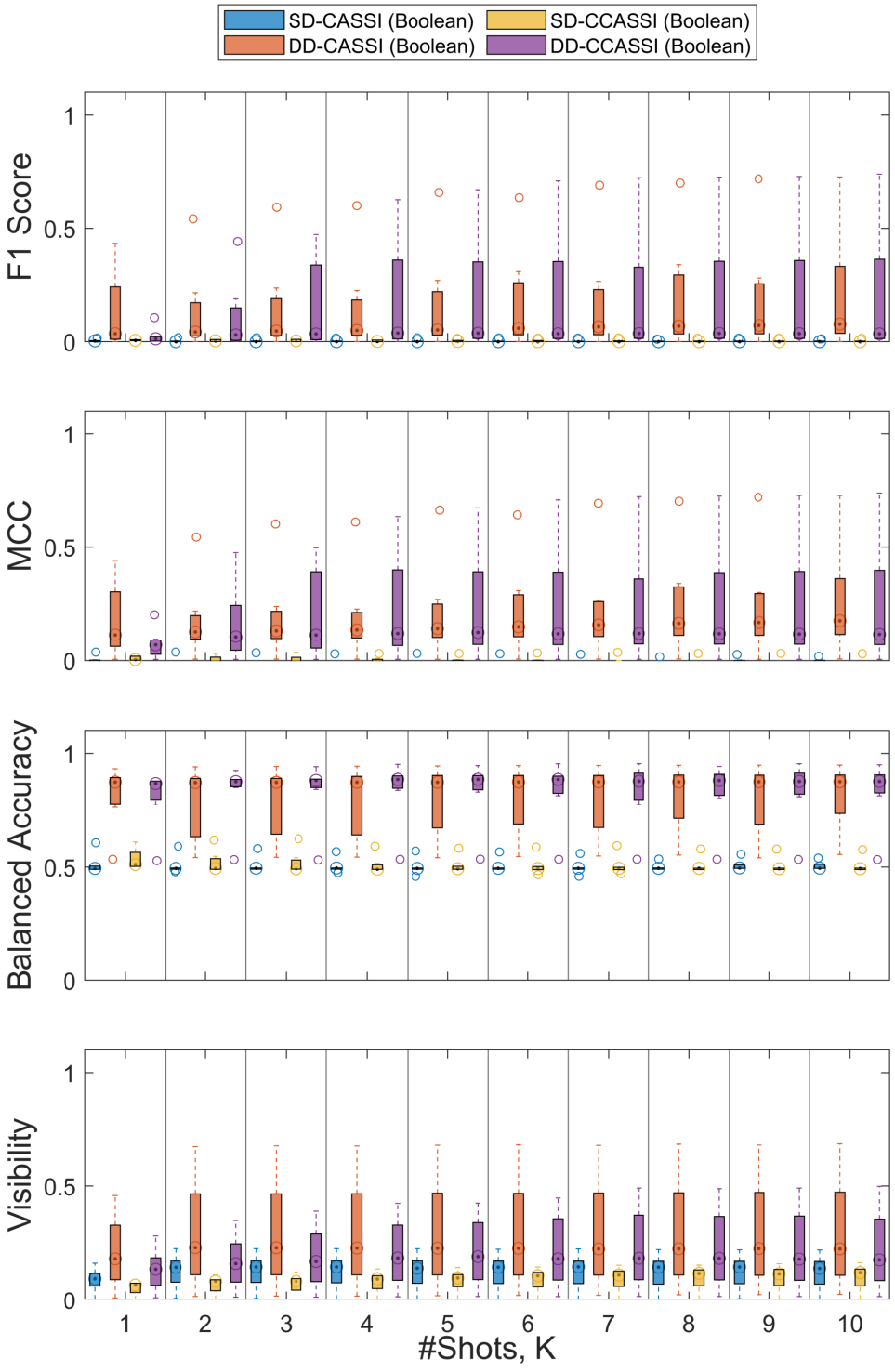


Figure 6.44: Per-shot comparison between binary and colour CASSI-TD approaches using a Boolean CA with a detection threshold of $T = 0.95$.

Table 6.8: CASSI-TD performance on an image from the UDRC dataset.

SD-CASSI (Boolean)															
$T = 0.9 - 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.61	0.01	0.28	0.02	0.04	0.56	0.01	0.15	0.01	0.02	0.54	0.00	0.11	0.01	0.01
2	0.49	0.00	0.63	0.00	0.00	0.48	0.00	0.46	0.00	0.00	0.49	0.00	0.34	0.00	0.00
3	0.49	0.00	0.01	0.00	0.00	0.49	0.00	0.00	0.00	0.00	0.49	0.00	0.00	0.00	0.00
4	0.51	0.00	0.05	0.01	0.00	0.50	0.00	0.01	0.00	0.00	0.50	0.00	0.00	0.00	0.00
5	0.49	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00
6	0.50	0.00	0.03	0.00	0.00	0.49	0.00	0.00	0.00	-0.01	0.49	0.00	0.00	0.00	-0.01
7	0.49	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00
Average	0.51	0.00	0.14	0.00	0.00	0.50	0.00	0.09	0.00	0.00	0.50	0.00	0.06	0.00	0.00
SD-CCASSI (Boolean)															
$T = 0.9 - 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.59	0.00	0.41	0.01	0.02	0.57	0.01	0.20	0.01	0.02	0.57	0.01	0.19	0.01	0.02
2	0.48	0.00	0.72	0.00	0.00	0.49	0.00	0.59	0.00	0.00	0.49	0.00	0.56	0.00	0.00
3	0.51	0.00	0.07	0.00	0.00	0.49	0.00	0.00	0.00	0.00	0.49	0.00	0.00	0.00	0.00
4	0.57	0.00	0.25	0.01	0.02	0.51	0.00	0.04	0.01	0.01	0.50	0.00	0.02	0.00	0.00
5	0.51	0.00	0.06	0.00	0.00	0.49	0.00	0.00	0.00	-0.01	0.50	0.00	0.00	0.00	0.00
6	0.54	0.00	0.14	0.01	0.01	0.50	0.00	0.02	0.00	0.00	0.49	0.00	0.00	0.00	-0.01
7	0.51	0.00	0.04	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00
Average	0.53	0.00	0.24	0.01	0.01	0.51	0.00	0.12	0.00	0.00	0.51	0.00	0.11	0.00	0.00
DD-CASSI (Boolean)															
$T = 0.9 - 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.82	0.01	0.83	0.01	0.06	0.84	0.01	0.74	0.03	0.08	0.84	0.01	0.73	0.02	0.08
2	0.53	0.00	0.87	0.00	0.00	0.57	0.00	0.83	0.00	0.01	0.61	0.00	0.83	0.00	0.01
3	0.86	0.03	0.74	0.05	0.13	0.85	0.11	0.71	0.14	0.21	0.85	0.18	0.72	0.19	0.26
4	0.90	0.02	0.85	0.03	0.11	0.91	0.18	0.83	0.20	0.29	0.91	0.31	0.83	0.35	0.42
5	0.85	0.03	0.72	0.05	0.13	0.85	0.11	0.72	0.13	0.21	0.85	0.18	0.71	0.19	0.27
6	0.75	0.17	0.50	0.23	0.27	0.73	0.77	0.46	0.51	0.55	0.76	0.80	0.53	0.58	0.61
7	0.72	0.34	0.44	0.34	0.36	0.66	0.28	0.32	0.27	0.29	0.70	0.34	0.39	0.32	0.34
Average	0.78	0.08	0.71	0.10	0.15	0.77	0.21	0.66	0.18	0.23	0.79	0.26	0.68	0.24	0.29
DD-CCASSI (Boolean)															
$T = 0.9 - 0.99$	K=1					K=5					K=10				
Target	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC	BAcc	P	R	F1	MCC
1	0.74	0.00	0.86	0.00	0.03	0.81	0.01	0.80	0.02	0.06	0.82	0.01	0.81	0.02	0.06
2	0.54	0.00	0.91	0.00	0.01	0.57	0.00	0.89	0.00	0.01	0.56	0.00	0.89	0.00	0.01
3	0.84	0.01	0.76	0.02	0.08	0.87	0.07	0.75	0.10	0.18	0.87	0.10	0.76	0.13	0.20
4	0.84	0.00	0.88	0.01	0.05	0.92	0.11	0.86	0.14	0.21	0.93	0.14	0.89	0.17	0.24
5	0.86	0.01	0.78	0.03	0.09	0.86	0.07	0.73	0.10	0.17	0.86	0.08	0.75	0.10	0.18
6	0.83	0.02	0.70	0.03	0.10	0.79	0.60	0.58	0.51	0.54	0.78	0.72	0.56	0.57	0.60
7	0.79	0.06	0.58	0.11	0.18	0.77	0.31	0.53	0.36	0.39	0.75	0.33	0.50	0.36	0.39
Average	0.78	0.02	0.78	0.03	0.07	0.80	0.17	0.74	0.17	0.22	0.80	0.20	0.74	0.19	0.24

BAcc - Balanced Accuracy, P - Precision, R - Recall, F1 - F1 Score, MCC - Matthews Correlation Coefficient.
1 - Brown Carpet, 2 - Green Carpet, 3 - Green Ceramic, 4 - Green Perspex, 5 - Grey Ceramic, 6 - Orange Perspex, 7 - White Perspex.

6.5 Summary

In this chapter, a novel approach for performing TD on compressed CASSI measurements, CASSI-TD has been presented. In addition to the CASSI-TD approach, models for simulating the behaviour of four CASSI architectures have been developed. Routines for producing the appropriate CAs and a database of CAs for obtaining CASSI measurements were created to generate reproducible results.

The binary and colour CASSI-TD models presented in this chapter have been shown to provide reasonable detections, with a selection of TD tasks on synthetic imagery, multispectral imagery from a controlled environment, as well as a remote sensing application displayed. In each case, the dual dispersion CASSI techniques were able to detect and localise spectra of interest consistently, although with varying specificity. The single dispersion CASSI-based TD was limited in its efficacy as such architectures distribute the energy of a single pixel in an input across multiple pixels in a compressed measurement, meaning they are less identifiable in the proposed approach. However, alternative TD algorithms may be developed for use with SD-CASSI-based approaches. Generally, areas of contiguous material were able to be detected, but the boundaries of objects, or indeed small objects, as highlighted with the remote sensing TD task, are missed when using single dispersion architectures within CASSI-TD.

Of the CA types tested, Boolean, VAC, and STBN CAs consistently outperformed the Random and PD CAs in each of the architectures tested. This was somewhat expected, due to each of these CA types including the constraint that each pixel of the image be sampled. As the performance of the CASSI-TD approach is evaluated on a per-pixel basis, this constraint is desirable as, by extension, it ensures that each pixel in a target will be sampled.

As the proposed CASSI-TD approach is based on a simple absolute difference between an observed image and the expected compressive measurement for a given spectrum and a known behaviour in a sensing matrix, $\boldsymbol{H}$, it is sensitive to variation in spectra. It was observed that similar spectra, and those with similar AUC, can produce similar compressive measurements using a CASSI architecture. Without any spectral information available to any CASSI-TD approach, the ability to suppress this is limited. Additionally, particularly dark spectra were difficult to isolate due to the relatively low energy and abundance of similar dark, albeit spectrally distinct, materials in a scene. Further investigation into the optimal design of colour CAs may assist in improving the spectral selectivity and visibility of targets in the compressed domain.

An additional appreciation is the number of elements within a sensing matrix, $\boldsymbol{H}$, based

on an ensemble CA. In a binary CA, that has been designed using one of the structured CA generation methods, each pixel has the constraint that it is sampled once. In the ensemble colour CA generation methods, this was changed such that each CA in the ensemble would sample the entire image, with the caveat that each element in the colour CA has a lower transmittance than an equivalent binary CA. This decision was made so as to not make the sensing matrix too sparse, however for larger values of K, the opposite holds true. In a binary CA-based sensing matrix, the maximum value in any row, and therefore the maximum value of the output, is equal to the number of image bands, L. However, in the current colour CA-based sensing matrices, this becomes dependent on the number of shots, K, as each pixel will be sampled by every CA in the ensemble. This results in a maximum value equal to $^{KL}/_{F}$, *i.e.,* the number of shots multiplied by the number of image bands divided by the number of filters, as the probability of each element appearing in a row of the sensing matrix is equal to $^{1}/_{F}$ – assuming the uniform spectral filters are being using to construct colour CAs. This was only discovered when taking CCASSI measurements and obtaining "brighter" compressed images than the equivalent binary CASSI architecture. An alternative would be to include the constraint that each pixel in an input image is sampled once by an ensemble colour CA, although this would have the effect of making the compressed measurements F times less bright and potentially lose information in the process.

There are numerous avenues for alternative and improved CASSI-TD algorithms. For example, one which takes into account spatial information in the compressed domain as opposed to a per-pixel approach as presented here. Alternatively, more robust similarity measures could replace the absolute difference measure used as the foundation of the TD approach presented in this chapter. The use of machine learning algorithms, have been used in the classification of CASSI measurements [212], therefore Mask-RCNN [257] or a similar approach, may be utilised for TD from compressive measurements to good effect, however, this would require substantial training data.

Despite these limitations, TD from compressed CASSI measurements has been demonstrated with a simple metric providing a promising benchmark for future investigation. Of the architectures simulated, DD-CASSI achieves the best performance in each of the datasets tested. Similarly to the reasoning from Chapter 5, additional results are not printed here to avoid saturation but are included in Appendices B.7 to B.9.

Chapter 7

Discussion, Conclusions and Future Work

7.1 Main Contributions and Findings

The main aims of this thesis were to investigate and develop novel signal and image processing methods for efficient and explainable object detection in spectral imagery. From this investigation three distinct areas of focus were identified.

1. A novel general extension to the HMT for multivariate data, the Multi-Dimensional Percentage Occupancy Hit-or-Miss Transform, has been developed and is presented in Chapter 4. This approach incorporates Percentage Occupancy-based noise robustness in order to detect objects in imagery under non-ideal circumstances and is suitable for use in binary, greyscale, colour, spectral, and volumetric imaging modalities. A set of design tools is presented for defining the optimal parameters in a given scenario. The proposed approach is applied to object detection in natural and synthetic single channel, colour, and hyperspectral imagery and outperforms similar single channel and multivariate morphological approaches. By performing pre-processing of an image to identify areas of interest, and applying the MDPOHMT in a subspace, both the efficacy and efficiency of the approach, when applied to hyperspectral TD, are improved.

2. Secondly, the Joint Spatio-Spectral Dimensionality Reduced Target Detection pipeline is presented in Chapter 5 and exploits the spatial and spectral redundancy in remote sensing HSI applications to improve the computational overhead of performing TD. It has been shown that applying spatial DR prior to spectral DR improves the performance of classical TD algorithms. Remote sensing indices have been shown to be beneficial in masking known non-target pixels, thereby improving the dimensionality reduced subspace in which TD is performed. This has the added benefit of further improving the compression rate of the data, with negligible impact on TD performance, with the metrics considered improving or being maintained with compression of $> 95\%$ achieved.

3. Finally, the novel CASSI-TD approach was developed, presented in Chapter 6, which seeks to exploit the inherent sparsity within spectral data using CSI. The ability to perform TD directly from compressed measurements of a scene without the need for reconstruction using CASSI-TD is shown. As the characteristics of the optical elements are known, so too is the behaviour of a desired target spectrum, enabling detection using relatively simple techniques. For remote sensing applications, due to the small object size, dual dispersion architectures are far more appropriate, however have a greater physical complexity. The design and use of both binary and colour CAs has been presented, and, in both cases, the use of Boolean CAs which ensure sufficient useful spatial and spectral context is encoded allows for reliable TD.

Each of these approaches can be used to improve TD in spectral imaging whilst reducing computational overhead by exploiting the inherent sparsity of MSI and HSI. The work presented here highlights the broad scope for improving upon TD techniques in high-dimensional, sparse, data; from the extension of classical image processing algorithms, to optimal compression techniques, and the simulation of imaging hardware. Each approach considered in this thesis, although from seemingly disparate domains of signal and image processing, has showcased a method of improving upon both the efficiency and efficacy of TD in spectral imaging.

7.2 Limitations of the Outlined Contributions

The contributions presented in this thesis, while having been developed extensively and presented fairly and accurately, are not without inherent limitations. These are discussed here, and solutions and potential areas of further investigation for each are covered in Section 7.3.

A) Chapter 4 - MDPOHMT

1. The addition of colour and spectral information to the designed Structuring Element used in the MDPOHMT, although a hallmark of the approach, is a departure from the relative simplicity behind the definition of binary and greyscale morphological operations. This complexity, while useful in pattern recognition and object detection tasks, requires more consideration when designing appropriate SEs. Additionally, by utilising multivariate SEs, the computational overhead of the approach is increased when compared to an equivalent single channel HMT.
2. The MDPOHMT, like other morphological operations is not scale, rotation, or translation invariant and therefore, when detecting complex, variable, or other

non-generalisable objects, multiple SEs must be created which can be a time intensive process and may require some expert knowledge. As discussed in Section 7.3 this may be an opportunity for some sort of automated tool for designing colour SEs given some initial parameters.

3. Similar to other filtering approaches, the current implementation of the MDPOHMT is inefficient when processing large images, or SEs, or both. As the size of an image or SE increases, the number of individual operations increases exponentially. While reasonable optimisation has been performed, this is still a limitation of the MDPOHMT in its current form and some form of fast implementation may improve this.

B) Chapter 5 - JSSDR-TD

1. As the datasets used to test the presented JSSDR-TD approach exhibited large quantities of vegetative background, a priority on utilising VIs in the spatial DR stage was necessary. This necessity, however, limits the approach in its current form to operation in similar scenarios. In order to translate the JSSDR-TD approach to alternative environments, an understanding of the features of these environments and how they can be simplified is required, and alternative measures for appropriate spatial DR need to be developed or incorporated.
2. The definition of the required number of components is an ill posed problem in DR and indeed may still be optimised further than the efforts shown in this approach. Additionally, some DR methods tested in this thesis require a specific number of input features, limiting the use of the JSSDR-TD pipeline in certain imagery without appropriate pre-processing prior to data ingestion.
3. The current spectral DR component of the JSSDR-TD pipeline requires the computation of an optimal reduced dimensionality representation of the imaged scene. This can incur significant computational overhead compared to the TD operations themselves and is a limiting factor in the overall performance of the proposed pipeline.

C) Chapter 6 - CASSI-TD

1. The simulated CASSI architectures presented in this thesis are simplistic, assuming ideal transmission of light through the optical components with no dispersion, noise, or interference modelled. As such, they may not be representative of a physical, hardware-based, system.

2. The absolute difference-based measure for comparing compressed measurements against an expected response given some target spectra is sufficient for a proof-of-concept as presented in this thesis. However, it has been shown, in places, to be sensitive to small variations in material, illumination, or noise in an image. Additionally, there is a limited amount of separability between similar colours or spectra with similar spectral densities. As a result, there is little overhead in thresholds for performing a binary classification between target and background. As such, in order to provide more robustness to these effects, alternative, more optimal, matching algorithms may be desirable.
3. The current approach utilises CA designs from the literature which are commonly used for obtaining optimal reconstructions. As this study focussed on performing TD, alternative definitions for binary and colour CAs may be more appropriate and provide improved performance.

7.3 Future Work

During the development of the techniques presented in this thesis, various areas of additional study which could form the basis of future work were identified. These are summarised below:

A) Chapter 4 - MDPOHMT

1. As discussed in this thesis, there are multiple methods for performing morphological operations, specifically the HMT in colour and multi-channel images. The use of rank order filters in the proposed MDPOHMT has been shown to aid in noise robustness of such operations. Combining alternative colour HMTs, such as the SHMT, with rank order filters and PO may produce improve results.
2. As well as other algorithms, there are a number of alternative similarity measures and colour spaces that represent colour information in various ways. The Euclidean distance measure and RGB colour space were considered in this thesis, as they are easily extensible to multispectral and hyperspectral images - a task which was identified as a key area to investigate. However, testing the MDPOHMT in alternative colour spaces may yield improvements as colours can be more easily differentiated.
3. Like with any other filter or convolution based image processing techniques, the performance of the MDPOHMT, and morphological operators more generally, degrades with increasing image and filter sizes. Investigations into mitigating this, such

as processing image pyramids or generating an ensemble of smaller SEs to perform the same morphological operation would reduce the computation time of the MDPOHMT and methods like it.

4. Feature pooling is a key step in CNN architectures between sets of convolutional layers. Features are commonly pooled by taking the maximum or average value within a window. This pooling has the effect of downsampling an input feature map and assists in extracting general information from an input by discarding fine detail. The use of rank order filters, where the optimal rank is a learned parameter in the backpropagation, or model training, process may assist in optimal feature extraction by allowing flexibility in the pooling stage.
5. Morphological filters have also been used as layers in CNN architectures, the HMT may have some use in performing feature extraction based on local regions in an input feature map.

B) Chapter 5 - JSSDR-TD

1. NDVI and similar VIs are used heavily in Chapter 5 to improve spectral DR and TD, this was due to the heavily vegetated areas present in the underlying image data. Whilst the presence of vegetative background and non-vegetative targets may be fairly ubiquitous in remote sensing and TD applications, other remote sensing indices may prove useful to perform similar spatial dimensionality reduction *i.e.,* the NDSI or similar. Alternatively, new methods for combining image bands to remove definite non-target areas may be an interesting development.
2. While a number of DR techniques are investigated and compared in this thesis, they all have one thing in common; the need to calculate optimal basis vectors in an *a priori* manner. Unlike known basis functions such as those in the DCT, this can be inefficient and complex and requires dedicated computation. A method of computing or learning an optimal compressed representation, such as autoencoders, for any input image would negate this requirement.
3. The spectral TD algorithms used in this thesis may also be used as similarity metrics in other techniques, for instance in an MDPOHMT performed on hyperspectral imagery. Application specific similarity measures may be used to provide improvements over more general distance measures.

C) Chapter 6 - CASSI-TD

1. As discussed in this thesis, the CA design is key to the operation of any CASSI system. However, as much of the current research focus surrounding CSI techniques includes the reconstruction problem, many of the CA design optimisations are made with this in mind. As a result, there may be alternative methods that produce optimal CAs for TD. Such CAs, and methods for producing them, may improve the CASSI-TD architecture presented in this thesis. Additionally, spatio-temporal colour CAs which constrain each element in a CCASSI system to be sampled once, and the resultant effect on compressed measurements may be evaluated.
2. A measure of how much information is retained in CASSI measurements, and therefore how much information about any potential targets contained within, would allow for optimisation of CAs for TD. Additionally, it would form a quantitative metric to use in training generative AI, or genetic algorithm-based, design tools for CA generation.
3. The TD algorithm presented as part of this thesis is incredibly simple. However, as the aim of this work was to investigate the feasibility of performing TD from compressed measurements and benchmark multiple simulated CASSI architectures in such tasks, improvements to the detection or matching algorithms may improve upon the results presented here. In particular, the use of machine learning to identify possible targets may be investigated.
4. The simulations presented in this thesis are highly simplified and idealistic. This was done so as to investigate whether performing TD from CASSI measurements was a possibility in such an ideal case. However, in order to verify that such techniques can be applied to hardware, more complex simulations accounting for optical interactions and noise are required.
5. With improved simulations and TD algorithms, an indication of how such techniques would respond to measurements captured using physical CASSI systems may be obtained. The behaviour of the hardware is known, much as the behaviour of a sensing matrix is known, however it is not an idealised system and as a result limitations not encountered in this thesis may be discovered.

References

[1] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis *et al.*, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris)," *Remote sensing of environment*, vol. 65, no. 3, pp. 227–248, 1998.

[2] L. J. Rickard, R. W. Basedow, E. F. Zalewski, P. R. Silverglate, and M. Landers, "Hydice: An airborne system for hyperspectral imaging," in *Imaging Spectrometry of the Terrestrial Environment*, vol. 1937. SPIE, 1993, pp. 173–179.

[3] D. McKeown and M. Richardson, "Improved airborne system for sensing wildfires," Rochester Institute of Technology, Tech. Rep., 2008.

[4] J. Aschbacher, "Esa's earth observation strategy and copernicus," *Satellite earth observations and their impact on society and policy*, pp. 81–86, 2017.

[5] M. A. Wulder, T. R. Loveland, D. P. Roy, C. J. Crawford, J. G. Masek, C. E. Woodcock, R. G. Allen, M. C. Anderson, A. S. Belward, W. B. Cohen *et al.*, "Current status of landsat program, science, and applications," *Remote sensing of environment*, vol. 225, pp. 127–147, 2019.

[6] P. Tatzer, M. Wolf, and T. Panner, "Industrial application for inline material sorting using hyperspectral imaging in the nir range," *Real-Time Imaging*, vol. 11, no. 2, pp. 99–107, 2005.

[7] D.-W. Sun, *Hyperspectral imaging for food quality analysis and control*. Elsevier, 2010.

[8] D. Saha and A. Manickavasagan, "Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review," *Current Research in Food Science*, vol. 4, pp. 28–44, 2021.

[9] L. M. Dale, A. Thewis, C. Boudry, I. Rotar, P. Dardenne, V. Baeten, and J. A. F. Pierna, "Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review," *Applied Spectroscopy Reviews*, vol. 48, no. 2, pp. 142–159, 2013.

[10] T. Adão, J. Hruška, L. Pádua, J. Bessa, E. Peres, R. Morais, and J. J. Sousa, "Hyperspectral imaging: A review on uav-based sensors, data processing and applications for agriculture and forestry," *Remote sensing*, vol. 9, no. 11, p. 1110, 2017.

[11] B. Lu, P. D. Dao, J. Liu, Y. He, and J. Shang, "Recent advances of hyperspectral imaging technology and applications in agriculture," *Remote Sensing*, vol. 12, no. 16, p. 2659, 2020.

[12] A. Richards, *Alien vision: Exploring the electromagnetic spectrum with Imaging Technology*, 2nd ed. SPIE Press, 2011.

[13] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: A review," *Ieee Access*, vol. 6, pp. 14 118–14 129, 2018.

[14] J. Serra, *Image Analysis and Mathematical Morphology.* New York: Academic press, 1982, vol. 1.

[15] ——, *Image Analysis and Mathematical Morphology, Vol. 2: Theoretical Advances.* New York, NY: Academic Press, 1988, vol. 2.

[16] M. L. Comer and E. J. Delp, "Morphological operations for color image processing," *J. Electronic Imaging*, vol. 8, no. 3, pp. 279–289, 1999.

[17] J. Serra, "The "false colour" problem," in *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing.* Springer, 2009, Conference Proceedings, pp. 13–23.

[18] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, no. 1, pp. 55–63, 1968.

[19] G. V. Trunk, "A problem of dimensionality: A simple example," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 3, pp. 306–307, 1979.

[20] C.-H. Chen and X. Zhang, "On the roles of pca and ica in data fusion," in *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No. 00CH37120)*, vol. 6. IEEE, 2000, pp. 2620–2622.

[21] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Kernel principal component analysis for feature reduction in hyperspectrale images analysis," in *Proceedings of the 7th Nordic Signal Processing Symposium - NORSIG 2006*, 2006, pp. 238–241.

[22] P. S. Thenkabail, M. K. Gumma, P. Teluguntla, and A. Irshad, "Hyperspectral remote sensing of vegetation and agricultural crops," *Photogrammetric Engineering & Remote Sensing (TSI)*, vol. 80, no. 8, pp. 695–723, 2014.

[23] G. Matheron, *Random sets and integral geometry.* New York, NY: Wiley, 1975.

[24] J. Serra, "Introduction to mathematical morphology," *Computer Vision, Graphics, and Image Processing*, vol. 35, no. 3, pp. 283–305, 1986.

[25] H. Heijmans, *Morphological Image Operators.* Academic Press, 1994.

[26] J. Serra, "The beginnings of cmm," 2017, une traduction en français est disponible en annexe Mathematics [math]/History and Overview [math.HO]Other publications. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01532876

[27] P. Soille, *Morphological image analysis: principles and applications*, 2nd ed. Springer Science & Business Media, 2013.

[28] A. N. Evans and X. U. Liu, "A morphological gradient approach to color edge detection," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1454–1463, 2006.

[29] H. Xu, Y. Zhang, and H. Zhao, "Edge detection of color image using mathematical morphology in hsv color space," in *Proceedings of 2012 2nd International Conference on Computer Science and Network Technology*, 2012, Conference Proceedings, pp. 2112–2116.

[30] N. R. Harvey, R. B. Porter, and J. Theiler, "Ship detection in satellite imagery using rank-order greyscale hit-or-miss transforms," Los Alamos National Laboratory (LANL), Report, 2010.

[31] N. Lanisa, N. S. Cheok, and L. K. Wee, "Color morphology and segmentation of the breast thermography image," in *2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, 2014, Conference Proceedings, pp. 772–775.

[32] E. Aptoula, S. Lefèvre, and C. Collet, "Mathematical morphology applied to the segmentation and classification of galaxies in multispectral images," pp. 1–5, 4-8 Sept. 2006 2006.

[33] G. Noyel, J. Angulo, and D. Jeulin, "Morphological segmentation of hyperspectral images," *Image Analysis and Stereology*, vol. 26, no. 3, pp. 101–109, 2007.

[34] B. Perret, S. Lefèvre, and C. Collet, "A robust hit-or-miss transform for template matching applied to very noisy astronomical images," *Pattern Recognition*, vol. 42, no. 11, pp. 2470–2480, 2009.

[35] S. R. Sternberg, "Grayscale morphology," *Computer Vision, Graphics, and Image Processing*, vol. 35, no. 3, pp. 333–355, 1986.

[36] P. Soille, "On morphological operators based on rank filters," *Pattern Recognition*, vol. 35, no. 2, pp. 527–535, 2002.

[37] C. Ronse, L. Najman, and E. Decencière, *Mathematical morphology: 40 years on: proceedings of the 7th International Symposium on Mathematical Morphology, April 18-20, 2005*. Springer Science & Business Media, 2006, vol. 30.

[38] J. Angulo, "Morphological colour operators in totally ordered lattices based on distances: Application to image filtering, enhancement and analysis," *Computer Vision and Image Understanding*, vol. 107, no. 1, pp. 56–73, 2007.

[39] S. Lefèvre, E. Aptoula, B. Perret, and J. Weber, *Morphological template matching in color images*. Springer, 2014, pp. 241–277.

[40] C. W. Yeh and D. Pycock, "Similarity-based colour morphology," in *2013 5th Computer Science and Electronic Engineering Conference (CEEC)*, 2013, Conference Proceedings, pp. 71–76.

[41] C. Ronse, "Why mathematical morphology needs complete lattices," *Signal Processing*, vol. 21, no. 2, pp. 129–154, 1990.

[42] B. Naegel, N. Passat, and C. Ronse, "Grey-level hit-or-miss transforms—part i: unified theory," *Pattern Recognition*, vol. 40, no. 2, pp. 635–647, 2007.

[43] M. Khosravi and R. W. Schafer, "Template matching based on a grayscale hit-or-miss transform," *IEEE Transactions on image Processing*, vol. 5, no. 6, pp. 1060–1066, 1996.

[44] C. Barat and C. Ducottet, "Pattern matching using morphological probing," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, vol. 1. IEEE, 2003, Conference Proceedings, pp. I–369.

[45] C. Ronse, "A lattice-theoretical morphological view on template extraction in images," *Journal of Visual Communication and Image Representation*, vol. 7, no. 3, pp. 273–295, 1996.

[46] P. Murray and S. Marshall, *A Review of Recent Advances in the Hit-or-Miss Transform.* Elsevier, 2013, vol. 175, book section 5, pp. 221–282.

[47] P. Maragos, "Optimal morphological approaches to image matching and object detection," in *[1988 Proceedings] Second International Conference on Computer Vision*, 1988, Conference Proceedings, pp. 695–699.

[48] Y.-H. Doh, J.-C. Kim, J.-W. Kim, S.-J. Kim, and M. S. Alam, "New morphological hit-miss transform-based detection algorithm," *Optical Engineering*, vol. 41, 2002.

[49] P. Murray, S. Marshall, and E. Bullinger, "The percentage occupancy hit or miss transform," in *Signal Processing Conference, 2009 17th European.* IEEE, 2009, Conference Proceedings, pp. 253–257.

[50] P. Murray and S. Marshall, "A new design tool for feature extraction in noisy images based on grayscale hit-or-miss transforms," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1938–1948, 2011.

[51] V. Barnett, "The ordering of multivariate data," *Journal of the Royal Statistical Society*, vol. 139, no. 3, pp. 318–355, 1976.

[52] E. Aptoula and S. Lefèvre, "A comparative study on multivariate mathematical morphology," *Pattern Recognition*, vol. 40, no. 11, pp. 2914–2929, 2007.

[53] M. Ivanovici, A. Căliman, N. Richard, and C. Fernandez-Maloigne, "Towards a multivariate probabilistic morphology for colour images," in *Conference on Colour in Graphics, Imaging, and Vision*, vol. 2012. Society for Imaging Science and Technology, 2012, Conference Proceedings, pp. 189–193.

[54] R.-M. Coliban and M. Ivanovici, "Color and multispectral texture characterization using pseudo-morphological tools," in *Image Processing (ICIP), 2014 IEEE International Conference on.* IEEE, 2014, Conference Proceedings, pp. 630–634.

[55] A. K. Katsaggelos, K.-T. Lay, and N. P. Galatsanos, "A general framework for frequency domain multi-channel signal processing," *IEEE Transactions on Image Processing*, vol. 2, no. 3, pp. 417–420, 1993.

[56] I. Pitas and P. Tsakalides, "Multivariate ordering in color image filtering," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 1, no. 3, pp. 247–259, 295–6, 1991.

[57] R. C. Hardie and G. R. Arce, "Ranking in $\mathbb{R}^p$ and its use in multivariate image estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 1, no. 2, pp. 197–209, 1991.

[58] A. Venetsanopoulos and P. Trahanias, "Vector directional filters-a new class of multichannel image processing filters," *IEEE Transactions on Image processing*, vol. 2, no. 4, 1993.

[59] J. Astola, P. Haavisto, and Y. Neuvo, "Vector median filters," *Proceedings of the IEEE*, vol. 78, no. 4, pp. 678–689, 1990.

[60] C. S. Regazzoni and A. Teschioni, "A new approach to vector median filtering based on space filling curves," *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 1025–1037, 1997.

[61] E. Stringa, A. Teschioni, and C. S. Regazzoni, "A classical morphological approach to color image filtering based on space filling curves," in *NSIP*, 1999, Conference Proceedings, pp. 351–354.

[62] P. Lambert and J. Chanussot, "Extending mathematical morphology to color image processing," *Proc. CGIP*, pp. 158–163, 2000.

[63] A. Ledda and W. Philips, "Majority ordering for colour mathematical morphology," in *2005 13th European Signal Processing Conference*, 2005, Conference Proceedings, pp. 1–4.

[64] H. M. Al-Otum, "Morphological operators for color image processing based on Mahalanobis distance measure," *Optical Engineering*, vol. 42, no. 9, pp. 2595 – 2606, 2003. [Online]. Available: https://doi.org/10.1117/1.1594727

[65] K. Stankov and D.-C. He, "Detection of buildings in multispectral very high spatial resolution images using the percentage occupancy hit-or-miss transform," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 10, pp. 4069–4080, 2014.

[66] J. Weber and S. Lefèvre, "A multivariate hit-or-miss transform for conjoint spatial and spectral template matching," in *Image and Signal Processing*. Springer Berlin Heidelberg, 2008, Conference Proceedings, pp. 226–235.

[67] ——, "Spatial and spectral morphological template matching," *Image and Vision Computing*, vol. 30, no. 12, pp. 934–945, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0262885612001047

[68] E. Aptoula, S. Lefèvre, and C. Ronse, "A hit-or-miss transform for multivariate images," *Pattern Recognition Letters*, vol. 30, no. 8, pp. 760–764, 2009.

[69] A. Ledoux, N. Richard, and A.-S. Capelle-Laizé, "Color hit-or-miss transform (cmomp)," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. IEEE, 2012, Conference Proceedings, pp. 2248–2252.

[70] A. Ledoux and N. Richard, "Morphology tools for color image processing, dermatology and cosmetology domain," in *2015 Conference on Cosmetic Measurements and Testing (COMET)*, 2015, Conference Proceedings, pp. 1–4.

[71] G. D. Finlayson, S. D. Hordley, and P. M. HubeL, "Color by correlation: a simple, unifying framework for color constancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1209–1221, 2001.

[72] A. Ledoux, N. Richard, and A. Capelle-Laizé, "How to specify or identify the most accurate multispectral distance function for mathematical morphology?" in *2013 Colour and Visual Computing Symposium (CVCS)*, 2013, Conference Proceedings, pp. 1–7.

[73] H. Deborah, N. Richard, and J. Y. Hardeberg, "A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 3224–3234, 2015.

[74] P. Bibiloni, M. González-Hidalgo, and S. Massanet, "Soft color morphology: A fuzzy approach for multivariate images," *Journal of Mathematical Imaging and Vision*, vol. 61, no. 3, pp. 394–410, 2019.

[75] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[76] K. Pearson, "On lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, vol. 2, no. 11, pp. 559–572, 1901.

[77] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of educational psychology*, vol. 24, no. 6, p. 417, 1933.

[78] J. Ren, J. Zabalza, S. Marshall, and J. Zheng, "Effective feature extraction and data reduction in remote sensing using hyperspectral imaging [applications corner]," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 149–154, 2014.

[79] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on geoscience and remote sensing*, vol. 26, no. 1, pp. 65–74, 1988.

[80] C. Gordon, "A generalization of the maximum noise fraction transform," *IEEE Transactions on geoscience and remote sensing*, vol. 38, no. 1, pp. 608–610, 2000.

[81] G. Luo, G. Chen, L. Tian, K. Qin, and S.-E. Qian, "Minimum noise fraction versus principal component analysis as a preprocessing step for hyperspectral imagery denoising," *Canadian Journal of Remote Sensing*, vol. 42, no. 2, pp. 106–116, 2016.

[82] J. B. Lee, A. S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 3, pp. 295–304, 1990.

[83] I. C. Chien and Q. Du, "Interference and noise-adjusted principal components analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5, pp. 2387–2396, 1999.

[84] C.-I. Chang, *Hyperspectral imaging: techniques for spectral detection and classification*. Springer Science & Business Media, 2003, vol. 1.

[85] J. Zabalza, J. Ren, M. Yang, Y. Zhang, J. Wang, S. Marshall, and J. Han, "Novel folded-pca for improved feature extraction and data reduction with hyperspectral imaging and sar in remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 112–122, 2014.

[86] Y. Jian, D. Zhang, A. F. Frangi, and Y. Jing-yu, "Two-dimensional pca: a new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.

[87] J. Xiuping and J. A. Richards, "Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 1, pp. 538–542, 1999.

[88] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[89] G. Franchi and J. Angulo, "Comparative study on morphological principal component analysis of hyperspectral images," in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2014 6th Workshop on*. IEEE, 2014, Conference Proceedings, pp. 1–4.

[90] ——, "Morphological principal component analysis for hyperspectral image analysis," *ISPRS International Journal of Geo-Information*, vol. 5, no. 6, p. 83, 2016.

[91] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4-5, pp. 411–430, 2000.

[92] L. J. Cao, K. S. Chua, W. K. Chong, H. P. Lee, and Q. M. Gu, "A comparison of pca, kpca and ica for dimensionality reduction in support vector machine," *Neurocomputing*, vol. 55, no. 1, pp. 321–336, 2003.

[93] Q. Du, I. Kopriva, and H. H. Szu, "Independent-component analysis for hyperspectral remote sensing imagery classification," *Optical Engineering*, vol. 45, no. 1, p. 017008, 2006.

[94] J. Wang and C. I. Chang, "Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1586–1600, 2006, 046pr Times Cited:191 Cited References Count:25.

[95] J. Wang and C. Chang, "Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 9, pp. 2601–2616, 2006.

[96] H. Wold, *Estimation of principal components and related models by iterative least squares*. Academic Press, NY, 1966, pp. 391–420.

[97] R. B. Cattell, "The scree test for the number of factors," *Multivariate Behavioral Research*, vol. 1, no. 2, pp. 245–276, 1966.

[98] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Transactions on geoscience and remote sensing*, vol. 42, no. 3, pp. 608–619, 2004.

[99] C. Yu, L. C. Lee, C. I. Chang, B. Xue, M. Song, and J. Chen, "Band-specified virtual dimensionality for band selection: An orthogonal subspace projection approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 5, pp. 2822–2832, 2018.

[100] R. Hadsell, S. Chopra, and Y. LeCun, "Dimensionality reduction by learning an invariant mapping," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 2. IEEE, 2006, Conference Proceedings, pp. 1735–1742.

[101] J. Zabalza, J. Ren, J. Zheng, H. Zhao, C. Qing, Z. Yang, P. Du, and S. Marshall, "Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging," *Neurocomputing*, vol. 185, pp. 1–10, 2016.

[102] J. Tschannerl, J. Ren, J. Zabalza, and S. Marshall, "Segmented autoencoders for unsupervised embedded hyperspectral band selection," in *2018 7th European Workshop on Visual Information Processing (EUVIP)*, 2018, Conference Proceedings, pp. 1–6.

[103] J. Tschannerl, J. Ren, P. Yuen, G. Sun, H. Zhao, Z. Yang, Z. Wang, and S. Marshall, "Mimr-dgsa: Unsupervised hyperspectral band selection based on information theory and a modified discrete gravitational search algorithm," *Information Fusion*, vol. 51, pp. 189–200, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253518300666

[104] J. Rouse Jr, R. Haas, J. Schell, and D. Deering, "Monitoring vegetation systems in the great plains with erts," in *Earth Resources Technology Satellite-1 Symposium, Washington DC*, vol. 3, 1974, Conference Proceedings, pp. 309–329.

[105] D. Haboudane, J. R. Miller, E. Pattey, P. J. Zarco-Tejada, and I. B. Strachan, "Hyperspectral vegetation indices and novel algorithms for predicting green lai of crop canopies: Modeling and validation in the context of precision agriculture," *Remote sensing of environment*, vol. 90, no. 3, pp. 337–352, 2004.

[106] A. K. Bhandari, A. Kumar, and G. K. Singh, "Feature extraction using normalized difference vegetation index (ndvi): A case study of jabalpur city," *Procedia Technology*, vol. 6, pp. 612–621, 2012.

[107] A. Young, A. Killey, S. Marshall, and A. Gray, "Outlier and target detection in aerial hyperspectral imagery," in *Hyperspectral Imaging and Applications Conference (HSI 2014)*, 2014, Conference Proceedings.

[108] A. Young, S. Marshall, and A. Gray, "Outlier and target detection in aerial hyperspectral imagery: a comparison of traditional and percentage occupancy hit or miss transform techniques," in *Automatic Target Recognition XXVI*, vol. 9844. International Society for Optics and Photonics, 2016, Conference Proceedings, p. 98440S.

[109] P. Hansen and J. Schjoerring, "Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression," *Remote sensing of environment*, vol. 86, no. 4, pp. 542–553, 2003.

[110] P. Ettehadi Osgouei, S. Kaya, E. Sertel, and U. Alganci, "Separating built-up areas from bare land in mediterranean cities using sentinel-2a imagery," *Remote Sensing*, vol. 11, no. 3, 2019.

[111] A. Gitelson and M. N. Merzlyak, "Spectral reflectance changes associated with autumn senescence of aesculus hippocastanum l. and acer platanoides l. leaves. spectral features and relation to chlorophyll estimation," *Journal of plant physiology*, vol. 143, no. 3, pp. 286–292, 1994.

[112] D. A. Sims and J. A. Gamon, "Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages," *Remote sensing of environment*, vol. 81, no. 2-3, pp. 337–354, 2002.

[113] S. K. McFeeters, "The use of the normalized difference water index (ndwi) in the delineation of open water features," *International Journal of Remote Sensing*, vol. 17, no. 7, pp. 1425–1432, 1996. [Online]. Available: https://doi.org/10.1080/01431169608948714

[114] H. Xu, "Modification of normalised difference water index (ndwi) to enhance open water features in remotely sensed imagery," *International Journal of Remote Sensing*, vol. 27, no. 14, pp. 3025–3033, 2006. [Online]. Available: https://doi.org/10.1080/01431160600589179

[115] B.-c. Gao, "Ndwi—a normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sensing of Environment*, vol. 58, no. 3, pp. 257–266, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0034425796000673

[116] D. K. Hall, G. A. Riggs, V. V. Salomonson, J. Barton, K. Casey, J. Chien, N. DiGirolamo, A. Klein, H. Powell, and A. Tait, "Algorithm theoretical basis document (atbd) for the modis snow and sea ice-mapping algorithms," *Nasa Gsfc*, vol. 45, 2001.

[117] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," *IEEE Transactions on geoscience and remote sensing*, vol. 32, no. 4, pp. 779–785, 1994.

[118] D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln laboratory journal*, vol. 14, no. 1, pp. 79–116, 2003.

[119] M. D. Farrell and R. M. Mersereau, "On the impact of pca dimension reduction for hyperspectral detection of difficult targets," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 2, pp. 192–195, 2005.

[120] X. Jin, S. Paswaters, and H. Cline, "A comparative study of target detection algorithms for hyperspectral imagery," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV*, vol. 7334. International Society for Optics and Photonics, 2009, Conference Proceedings, p. 73341W.

[121] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, "Is there a best hyperspectral detection algorithm?" in *Algorithms and technologies for multispectral, hyperspectral, and ultraspectral imagery XV*, vol. 7334. International Society for Optics and Photonics, 2009, Conference Proceedings, p. 733402.

[122] C. Olson, J. Nichols, J. Michalowicz, and F. Bucholtz, *Improved outlier identification in hyperspectral imaging via nonlinear dimensionality reduction*, ser. SPIE Defense, Security, and Sensing. SPIE, 2010, vol. 7695. [Online]. Available: https://doi.org/10.1117/12.851811

[123] D. Manolakis, M. Pieper, E. Truslow, T. Cooley, M. Brueggeman, and S. Lipson, "The remarkable success of adaptive cosine estimator in hyperspectral target detection," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX*, vol. 8743. International Society for Optics and Photonics, 2013, Conference Proceedings, p. 874302.

[124] S. Bakken, M. Orlandic, and T. A. Johansen, "The effect of dimensionality reduction on signature-based target detection for hyperspectral remote sensing," in *CubeSats and SmallSats for Remote Sensing III*, vol. 11131. International Society for Optics and Photonics, 2019, Conference Proceedings, p. 111310L.

[125] F. Macfarlane, P. Murray, S. Marshall, and H. White, "A fast hyperspectral hit-or-miss transform with integrated projection-based dimensionality reduction," in *Hyperspectral Imaging HSI*, 2018, Conference Proceedings.

[126] ——, *Object detection and classification in aerial hyperspectral imagery using a multivariate hit-or-miss transform*, ser. SPIE Defense + Commercial Sensing. SPIE, 2019, vol. 10986. [Online]. Available: https://doi.org/10.1117/12.2518103

[127] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, and M. Brueggeman, "Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 24–33, 2013.

[128] E. Ghaderpour and T. Vujadinovic, "Change detection within remotely sensed satellite image time series via spectral analysis," *Remote Sensing*, vol. 12, no. 23, p. 4001, 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/23/4001

[129] N. M. Nasrabadi, "Hyperspectral target detection: An overview of current and future challenges," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 34–44, 2014.

[130] P. Bajorski, "Analytical comparison of the matched filter and orthogonal subspace projection detectors for hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 7, pp. 2394–2402, 2007.

[131] W. Gross, J. Boehler, H. Schilling, W. Middelmann, J. Weyermann, P. Wellig, R. Oechslin, and M. Kneubühler, "Assessment of target detection limits in hyperspectral data," in *Target and Background Signatures*, vol. 9653. International Society for Optics and Photonics, 2015, Conference Proceedings, p. 96530J.

[132] L. L. Scharf and L. T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Conference Record of the Thirtieth Asilomar Conference on Signals, Systems and Computers*. IEEE, 1996, Conference Proceedings, pp. 1114–1117.

[133] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: a uniformly most-powerful-invariant adaptive detection statistic," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 427–438, 2005.

[134] M. Pieper, D. Manolakis, R. Lockwood, T. Cooley, P. Armstrong, and J. Jacobson, "Hyperspectral detection and discrimination using the ace algorithm," in *Imaging Spectrometry XVI*, vol. 8158. International Society for Optics and Photonics, 2011, Conference Proceedings, p. 815807.

[135] B. Alvey, A. Zare, M. Cook, and D. K. Ho, "Adaptive coherence estimator (ace) for explosive hazard detection using wideband electromagnetic induction (wemi)," in *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXI*, vol. 9823. International Society for Optics and Photonics, 2016, Conference Proceedings, p. 982309.

[136] E. Truslow, "Performance evaluation of the adaptive cosine estimator detector for hyperspectral imaging applications," Thesis, Northeastern University, 2012.

[137] A. Ziemann and J. Theiler, "Simplex ace: a constrained subspace detector," *Optical Engineering*, vol. 56, no. 8, p. 081808, 2017.

[138] K. Jnawali, J. P. Kerekes, and N. Rao, "Comparative study of spectral matched filter, constrained energy minimization, and adaptive coherence estimator for subpixel target detection based on hyperspectral imaging," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIV*, vol. 10644. International Society for Optics and Photonics, 2018, Conference Proceedings, p. 106441V.

[139] F. A. Kruse, A. Lefkoff, J. Boardman, K. Heidebrecht, A. Shapiro, P. Barloon, and A. Goetz, "The spectral image processing system (sips)—interactive visualization and analysis of imaging spectrometer data," *Remote sensing of environment*, vol. 44, no. 2-3, pp. 145–163, 1993.

[140] C.-I. Chang, "Spectral information divergence for hyperspectral image analysis," in *IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No. 99CH36293)*, vol. 1. IEEE, 1999, Conference Proceedings, pp. 509–511.

[141] Y. Du, C.-I. Chang, H. Ren, C.-C. Chang, J. O. Jensen, and F. M. D'Amico, "New hyperspectral discrimination measure for spectral characterization," *Optical Engineering*, vol. 43, no. 8, pp. 1777–1787, 2004.

[142] C.-I. Chang, H. Cao, and M. Song, "Orthogonal subspace projection target detector for hyperspectral anomaly detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4915–4932, 2021.

[143] S. Matteoli, F. Carnesecchi, M. Diani, G. Corsini, and L. Chiarantini, *Comparative analysis of hyperspectral anomaly detection strategies on a new high spatial and spectral resolution data set*, ser. SPIE Remote Sensing. SPIE, 2007, vol. 6748. [Online]. Available: https://doi.org/10.1117/12.738062

[144] C.-I. Chang, *Hyperspectral imaging: techniques for spectral detection and classification*. Springer Science & Business Media, 2003, vol. 1.

[145] ——, "Orthogonal subspace projection (osp) revisited: A comprehensive study and analysis," *IEEE transactions on geoscience and remote sensing*, vol. 43, no. 3, pp. 502–518, 2005.

[146] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, no. 7, pp. 5–28, 2010.

[147] M. Winter, *N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data*, ser. SPIE's International Symposium on Optical Science, Engineering, and Instrumentation. SPIE, 1999, vol. 3753. [Online]. Available: https://doi.org/10.1117/12.366289

[148] J. H. Gruninger, A. J. Ratkowski, and M. L. Hoke, "The sequential maximum angle convex cone (smacc) endmember model," in *Algorithms and technologies for multispectral, hyperspectral, and ultraspectral imagery X*, vol. 5425. International Society for Optics and Photonics, 2004, Conference Proceedings, pp. 1–15.

[149] F. Chaudhry, C.-C. Wu, W. Liu, C.-I. Chang, and A. J. Plaza, "Pixel purity index-based algorithms for endmember extraction from hyperspectral imagery," in *Recent Advances in Hyperspectral Signal and Image Processing*, 2006. [Online]. Available: https://api.semanticscholar.org/CorpusID:35969311

[150] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005.

[151] I. S. Reed and X. Yu, "Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1760–1770, 1990.

[152] F. Verdoja and M. Grangetto, "Graph laplacian for image anomaly detection," *Machine Vision and Applications*, vol. 31, no. 1, pp. 1–16, 2020.

[153] C.-I. Chang and S.-S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 6, pp. 1314–1325, 2002.

[154] D. Borghys, I. Kåsen, V. Achard, and C. Perneel, "Hyperspectral anomaly detection: Comparative evaluation in scenes with diverse complexity," *Journal of Electrical and Computer Engineering*, vol. 2012, p. 5, 2012.

[155] A. Chowdhury and M. Alam, *Target detection in hyperspectral imagery using noise adjusted principal component analysis and orthogonal subspace projection*, ser. SPIE Remote Sensing. SPIE, 2006, vol. 6365. [Online]. Available: https://doi.org/10.1117/12.684103

[156] S. Bakken, "Dimensionality reduction and target detection in hyperspectral remote sensing," Thesis, NTNU, 2018.

[157] R. Baraniuk, M. A. Davenport, M. F. Duarte, and C. Hegde, "An introduction to compressive sensing," *Connexions e-textbook*, pp. 24–76, 2011.

[158] R. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 08/08/07 2007.

[159] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge: Cambridge University Press, 2012.

[160] E. J. Candès, "Compressive sampling," in *Proceedings of the international congress of mathematicians*, vol. 3. Madrid, Spain, 2006, Conference Proceedings, pp. 1433–1452.

[161] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE transactions on information theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[162] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006, cited By :3481 Export Date: 18 October 2017.

[163] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006, cited By :8271 Export Date: 18 October 2017.

[164] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2004.

[165] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[166] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2992–3004, 2007.

[167] S. P. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, pp. 1–122, 2011. [Online]. Available: https://api.semanticscholar.org/CorpusID:51789432

[168] Q. Yue, B. Chen, Z. Xu, X. Liu, Y. Zhang, and Z. Zheng, "Deep-learning-based CASSI reconstruction with optical aberration correction," in *Optoelectronic Imaging and Multimedia Technology X*, Q. Dai, T. Shimura, and Z. Zheng, Eds., vol. 12767, International Society for Optics and Photonics. SPIE, 2023, p. 1276711. [Online]. Available: https://doi.org/10.1117/12.2686668

[169] I. Ahmed, A. Khalil, I. Ahmed, and J. Frnda, "Sparse signal representation, sampling, and recovery in compressive sensing frameworks," *IEEE Access*, vol. 10, pp. 85 002–85 018, 2022.

[170] J. Bacca, E. Martinez, and H. Arguello, "Computational spectral imaging: a contemporary overview," *J. Opt. Soc. Am. A*, vol. 40, no. 4, pp. C115–C125, Apr 2023. [Online]. Available: https://opg.optica.org/josaa/abstract.cfm?URI=josaa-40-4-C115

[171] W. Sun, C. Liu, J. Li, Y. M. Lai, and W. Li, "Low-rank and sparse matrix decomposition-based anomaly detection for hyperspectral imagery," *Journal of Applied Remote Sensing*, vol. 8, no. 1, 2014.

[172] A. W. Bitar, L.-F. Cheong, and J.-P. Ovarlez, "Sparse and low-rank matrix decomposition for automatic target detection in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, 2019.

[173] W. Li, M. Meng, and C. Zhao, "Hyperspectral target detection based on sparse errors matrix," in *Progress in Electromagnetics Research Symposium-Fall (PIERS-FALL), 2017*. IEEE, 2017, Conference Proceedings, pp. 3058–3062.

[174] L. Zhang and C. Zhao, "Hyperspectral anomaly detection based on spectral–spatial background joint sparse representation," *European Journal of Remote Sensing*, vol. 50, no. 1, pp. 362–376, 2017.

[175] R. Otazo, E. Candès, and D. K. Sodickson, "Low-rank plus sparse matrix decomposition for accelerated dynamic mri with separation of background and dynamic components," *Magnetic Resonance in Medicine*, vol. 73, no. 3, pp. 1125–1136, 2015.

[176] T. Zhou and D. Tao, "Godec: Randomized low-rank & sparse matrix decomposition in noisy case," in *Proceedings of the 28th International Conference on Machine Learning, ICML 2011*, 2011, Conference Proceedings.

[177] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, 2008.

[178] G. M. Gibson, S. D. Johnson, and M. J. Padgett, "Single-pixel imaging 12 years on: a review," *Optics Express*, vol. 28, no. 19, pp. 28 190–28 208, 2020. [Online]. Available: http://www.osapublishing.org/oe/abstract.cfm?URI=oe-28-19-28190

[179] B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. J. Padgett, "3d computational imaging with single-pixel detectors," *Science*, vol. 340, no. 6134, pp. 844–847, 2013.

[180] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4053–4085, 2011.

[181] C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," *Scientific Reports*, vol. 8, no. 1, p. 2369, 2018. [Online]. Available: https://doi.org/10.1038/s41598-018-20521-y

[182] H. Garcia, C. V. Correa, and H. Arguello, "Optimized sensing matrix for single pixel multi-resolution compressive spectral imaging," *IEEE Transactions on Image Processing*, vol. 29, pp. 4243–4253, 2020.

[183] A. Jerez, H. Garcia, and H. Arguello, "Single pixel spectral image fusion with side information from a grayscale sensor," in *2018 IEEE 1st Colombian Conference on Applications in Computational Intelligence (ColCACI)*, 2018, Conference Proceedings, pp. 1–6.

[184] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 105–115, 2014.

[185] G. R. Arce, H. Rueda, C. V. Correa, A. Ramirez, and H. Arguello, "Snapshot compressive multispectral cameras," *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–22, 2017. [Online]. Available: https://doi.org/10.1002/047134608X.W8345

[186] H. F. Rueda Chacón and H. Arguello Fuentes, "Spatial super-resolution in coded aperture-based optical compressive hyperspectral imaging systems," *Revista Facultad de Ingeniería Universidad de Antioquia*, no. 67, pp. 7–18, 2013.

[187] H. F. Rueda, H. Arguello, and G. R. Arce, "On super-resolved coded aperture spectral imaging," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX*, vol. 8743. International Society for Optics and Photonics, 2013, Conference Proceedings, p. 87431B.

[188] A. Ramirez, H. Arguello, G. R. Arce, and B. M. Sadler, "Spectral image classification from optimal coded-aperture compressive measurements," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3299–3309, 2013.

[189] J. M. Ramirez and H. Arguello, "Spectral image classification from multi-sensor compressive measurements," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 1, pp. 626–636, 2019.

[190] M. Busuioceanu, D. W. Messinger, J. B. Greer, and J. C. Flake, "Evaluation of the cassi-dd hyperspectral compressive sensing imaging system," in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX*, vol. 8743. International Society for Optics and Photonics, 2013, Conference Proceedings, p. 87431V.

[191] D. A. Boada Supelano, H. M. Vargas García, J. O. Albarracín Ferreira, and H. Argüello Fuentes, "A sparsity-based approach for spectral image target detection from compressive measurements acquired by the cassi architecture," *Ingeniería y Universidad*, vol. 21, no. 2, pp. 273–288, 2017.

[192] H. Arguello, A. Parada, and G. R. Arce, "Optimization of pseudorandom coded apertures for compressive spectral imaging," in *Compressive Sensing II*, vol. 8717. International Society for Optics and Photonics, 2013, Conference Proceedings, p. 87170D.

[193] C. V. Correa, H. Arguello, and G. R. Arce, "Spatiotemporal blue noise coded aperture design for multi-shot compressive spectral imaging," *Journal of the Optical Society of America A*, vol. 33, no. 12, pp. 2312–2322, 2016, http://momentsingraphics.de/BlueNoise.html. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-33-12-2312

[194] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied optics*, vol. 47, no. 10, pp. B44–B51, 2008.

[195] L. Galvis, D. Lau, X. Ma, H. Arguello, and G. R. Arce, "Coded aperture design in compressive spectral imaging based on side information," *Applied Optics*, vol. 56, no. 22, pp. 6332–6340, 2017. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-56-22-6332

[196] C. V. Correa, C. A. Hinojosa, G. R. Arce, and H. Arguello, "Multiple snapshot colored compressive spectral imager," *Optical Engineering*, vol. 56, no. 4, p. 041309, 2016.

[197] A. Parada-Mayorga, D. L. Lau, J. H. Giraldo, and G. R. Arce, "Blue-noise sampling on graphs," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 5, no. 3, pp. 554–569, 2019.

[198] J. Bacca, L. Galvis, and H. Arguello, "Coupled deep learning coded aperture design for compressive image classification," *Optics Express*, vol. 28, no. 6, pp. 8528–8540, 2020. [Online]. Available: http://www.osapublishing.org/oe/abstract.cfm?URI=oe-28-6-8528

[199] H. Arguello and G. R. Arce, "Colored coded aperture design by concentration of measure in compressive spectral imaging," *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1896–1908, 2014.

[200] H. Arguello, Y. Mejia, and G. Arce, "Colored coded apertures optimization in compressive spectral imaging by restricted isometry property," in *2014 IEEE International Conference on Image Processing (ICIP)*, 2014, Conference Proceedings, pp. 600–604.

[201] C. V. Correa, H. Arguello, and G. R. Arce, "Snapshot colored compressive spectral imager," *Journal of the Optical Society of America A*, vol. 32, no. 10, pp. 1754–1763, 2015. [Online]. Available: http://www.osapublishing.org/josaa/abstract.cfm?URI=josaa-32-10-1754

[202] A. Parada-Mayorga and G. R. Arce, "Colored coded aperture design in compressive spectral imaging via minimum coherence," *IEEE Transactions on Computational Imaging*, vol. 3, no. 2, pp. 202–216, 2017.

[203] C. H. F. Rueda, G. A. R. Calderón, and H. A. Fuentes, "Spectral selectivity in compressive spectral imaging based on grayscale coded apertures," in *Symposium of Signals, Images and Artificial Vision - 2013: STSIVA - 2013*, 2013, Conference Proceedings, pp. 1–5.

[204] X. Lin, Y. Liu, J. Wu, and Q. Dai, "Spatial-spectral encoded compressive hyperspectral imaging," *ACM Trans. Graph.*, vol. 33, no. 6, p. Article 233, 2014. [Online]. Available: https://doi.org/10.1145/2661229.2661262

[205] X. Lin, G. Wetzstein, Y. Liu, and Q. Dai, "Dual-coded compressive hyperspectral imaging," *Optics letters*, vol. 39, no. 7, pp. 2044–2047, 2014.

[206] J. Bacca, C. V. Correa, and H. Arguello, "A non-iterative reconstruction algorithm for single pixel spectral imaging with side information," in *2019 27th European Signal Processing Conference (EUSIPCO)*. IEEE, 2019, Conference Proceedings, pp. 1–5.

[207] L. Galvis, H. Arguello, D. Lau, and G. Arce, *Side information in coded aperture compressive spectral imaging*, ser. SPIE OPTO. SPIE, 2017, vol. 10117. [Online]. Available: https://doi.org/10.1117/12.2252167

[208] Y. Oiknine, I. August, V. Farber, D. Gedalin, and A. Stern, "Compressive sensing hyperspectral imaging by spectral multiplexing with liquid crystal," *Journal of Imaging*, vol. 5, no. 1, p. 3, 2019.

[209] Y. Oiknine, D. Gedalin, I. August, D. G. Blumberg, S. R. Rotman, and A. Stern, "Target detection with compressive sensing hyperspectral images," in *Image and Signal Processing for Remote Sensing XXIII*, vol. 10427. International Society for Optics and Photonics, 2017, Conference Proceedings, p. 104270O.

[210] D. Gedalin, Y. Oiknine, I. August, D. G. Blumberg, S. R. Rotman, and A. Stern, "Performance of target detection algorithm in compressive sensing miniature ultraspectral imaging compressed sensing system," *Optical Engineering*, vol. 56, no. 4, p. 041312, 2017.

[211] N. Cohen, S. Shmilovich, Y. Oiknine, and A. Stern, "Deep neural network classification in the compressively sensed spectral image domain," *Journal of Electronic Imaging*, vol. 30, no. 4, p. 041406, 2021.

[212] H. Zhang, X. Ma, X. Zhao, and G. R. Arce, "Compressive hyperspectral image classification using a 3d coded convolutional neural network," *Opt. Express*, vol. 29, no. 21, pp. 32 875–32 891, Oct 2021. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-29-21-32875

[213] N. Bouaynaya, M. Charif-Chefchaouni, and D. Schonfeld, "Theoretical foundations of spatially-variant mathematical morphology part i: Binary images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 5, pp. 823–836, 2008.

[214] B. Naegel, N. Passat, and C. Ronse, "Grey-level hit-or-miss transforms—part ii: Application to angiographic image processing," *Pattern Recognition*, vol. 40, no. 2, pp. 648–658, 2007.

[215] E. Aptoula and S. Lefèvre, "Pseudo multivariate morphological operators based on α-trimmed lexicographical extrema," in *2007 5th International Symposium on Image and Signal Processing and Analysis*. IEEE, 2007, Conference Proceedings, pp. 367–372.

[216] ——, "On lexicographical ordering in multivariate mathematical morphology," *Pattern Recognition Letters*, vol. 29, no. 2, pp. 109–118, 2008.

[217] C.-W. Yeh, "Colour morphology and its approaches," Thesis, University of Birmingham, 2015.

[218] A. F. Goetz, "Three decades of hyperspectral remote sensing of the earth: A personal view," *Remote Sensing of Environment*, vol. 113, pp. S5–S16, 2009, imaging Spectroscopy Special Issue. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S003442570900073X

[219] X. Dong, X. Xiao, Y. Pan, G. Wang, and Y. Yu, "Dmd-based hyperspectral imaging system with tunable spatial and spectral resolution," *Opt. Express*, vol. 27, no. 12, pp. 16 995–17 006, Jun 2019. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-27-12-16995

[220] S. Grusche, "Basic slit spectroscope reveals three-dimensional scenes through diagonal slices of hyperspectral cubes," *Appl. Opt.*, vol. 53, no. 20, pp. 4594–4603, Jul 2014. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-53-20-4594

[221] S. Song, D. Gibson, S. Ahmadzadeh, H. O. Chu, B. Warden, R. Overend, F. Macfarlane, P. Murray, S. Marshall, M. Aitkenhead, D. Bienkowski, and R. Allison, "Low-cost hyper-spectral imaging system using a linear variable bandpass filter for agritech applications," *Applied Optics*, vol. 59, no. 5, pp. A167–A175, 2020. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-59-5-A167

[222] G. Vane, R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen, and W. M. Porter, "The airborne visible/infrared imaging spectrometer (aviris)," *Remote Sensing of Environment*, vol. 44, no. 2, pp. 127–143, 1993, airbone Imaging Spectrometry. [Online]. Available: https://www.sciencedirect.com/science/article/pii/003442579390012M

[223] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural computation*, vol. 9, no. 7, pp. 1483–1492, 1997.

[224] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-gaussian signals," in *IEE proceedings F (radar and signal processing)*, vol. 140. IET, 1993, Conference Proceedings, pp. 362–370.

[225] C. Chang, "A review of virtual dimensionality for hyperspectral imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 4, pp. 1285–1305, 2018.

[226] A. Martinez-Uso, F. Pla, J. M. Sotoca, and P. García-Sevilla, "Clustering-based hyperspectral band selection using information measures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 4158–4171, 2007.

[227] A. Martínez-Usó, F. Pla, J. M. Sotoca, and P. García-Sevilla, "Comparison of unsupervised band selection methods for hyperspectral imaging," in *Pattern Recognition and Image Analysis: Third Iberian Conference, IbPRIA 2007, Girona, Spain, June 6-8, 2007, Proceedings, Part I 3*. Springer, 2007, pp. 30–38.

[228] J. Rouse Jr, R. Haas, D. Deering, J. Schell, and J. Harlan, "Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation," Remote Sensing Centre, Texas A&M University, Report, 1974.

[229] J. Xue and B. Su, "Significant remote sensing vegetation indices: A review of developments and applications," *Journal of Sensors*, vol. 2017, 2017.

[230] D. Qian, R. Hsuan, and I. C. Chein, "A comparative study for orthogonal subspace projection and constrained energy minimization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1525–1529, 2003.

[231] S. S. Shen, P. Bajorski, and P. E. Lewis, "Impact of missing endmembers on the performance of the osp detector for hyperspectral images," 2007.

[232] P. Bajorski, "Target detection under misspecified models in hyperspectral images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 470–477, 2012.

[233] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (roc) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

[234] J. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 251–255, 2008.

[235] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets," *PloS one*, vol. 10, no. 3, p. e0118432, 2015.

[236] D. M. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.

[237] C. M. Bishop, *Pattern Recognition and Machine Learning*, ser. Information Science and Statistics. New York, NY: Springer Science+Business Media, 2007.

[238] B. W. Matthews, "Comparison of the predicted and observed secondary structure of t4 phage lysozyme," *Biochimica et Biophysica Acta (BBA) - Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0005279575901099

[239] K. H. Brodersen, C. S. Ong, K. E. Stephan, and J. M. Buhmann, "The balanced accuracy and its posterior distribution," in *2010 20th international conference on pattern recognition*. IEEE, 2010, Conference Proceedings, pp. 3121–3124.

[240] D. Hand and P. Christen, "A note on using the f-measure for evaluating record linkage algorithms," *Statistics and Computing*, vol. 28, no. 3, pp. 539–547, 2017.

[241] D. Chicco and G. Jurman, "The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation," *BMC genomics*, vol. 21, no. 1, pp. 1–13, 2020.

[242] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[243] M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Optics express*, vol. 15, no. 21, pp. 14 013–14 027, 2007.

[244] N. Diaz, C. Hinojosa, and H. Arguello, "Adaptive grayscale compressive spectral imaging using optimal blue noise coding patterns," *Optics & Laser Technology*, vol. 117, pp. 147–157, 2019, https://github.com/carlosh93/Adaptive-Grayscale-CSI-Blue-Noise-Patterns. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0030399218312088

[245] H. Zhang, X. Ma, and G. R. Arce, "Compressive spectral imaging based on hexagonal blue noise coded apertures," *arXiv preprint arXiv:1905.09387*, 2019.

[246] D. Zhang, Q. Wang, F. Lin, X. Yin, C. Gu, and H. Qiao, "Development and evaluation of a new spectral disease index to detect wheat fusarium head blight using hyperspectral imaging," *Sensors (Basel)*, vol. 20, no. 8, 2020, zhang, Dongyan Wang, Qian Lin, Fenfang Yin, Xun Gu, Chunyan Qiao, Hongbo eng 18030701209/Anhui Provincial Major Science and Technology Projects 41771463/National Natural Science Foundation of China 41771469/National Natural Science Foundation of China Switzerland Sensors (Basel). 2020 Apr 16;20(8). pii: s20082260. doi: 10.3390/s20082260. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/32316216

[247] R. Bridson, "Fast poisson disk sampling in arbitrary dimensions." *SIGGRAPH sketches*, vol. 10, no. 1, p. 1, 2007.

[248] A. Wolfe, "Using blue noise for ray traced soft shadows," *Ray Tracing Gems II: Next Generation Real-Time Rendering with DXR, Vulkan, and OptiX*, pp. 367–394, 2021.

[249] R. A. Ulichney, *Void-and-cluster method for dither array generation*, ser. IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology. SPIE, 1993, vol. 1913. [Online]. Available: https://doi.org/10.1117/12.152707

[250] D. P. Mitchell, "Spectrally optimal sampling for distribution ray tracing," in *Proceedings of the 18th annual conference on Computer graphics and interactive techniques*, 1991, pp. 157–164.

[251] S. Wu and Y. Xu, "Dsn: A new deformable subnetwork for object detection," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2019.

[252] L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, and R. Qu, "A survey of deep learning-based object detection," *IEEE Access*, vol. 7, pp. 128 837–128 868, 2019.

[253] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, Conference Proceedings, pp. 91–99.

[254] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, Conference Proceedings, pp. 580–587.

[255] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, Conference Proceedings, pp. 1440–1448.

[256] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," July 01, 2018 2018. [Online]. Available: https://ui.adsabs.harvard.edu/#abs/2018arXiv180705511Z

[257] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, Conference Proceedings, pp. 2980–2988.

[258] R. P. Loce and E. R. Dougherty, "Facilitation of optimal binary morphological filter design via structuring element libraries and design constraints," *Optical Engineering*, vol. 31, no. 5, pp. 1008–1025, 18, 1992. [Online]. Available: https://doi.org/10.1117/12.56144

[259] S. Marshall, N. R. Harvey, and D. Greenhalgh, "Design of morphological filters using genetic algorithms," in *Signal Processing Conference, 2000 10th European.* IEEE, 2000, Conference Proceedings, pp. 1–4.

[260] K. Nogueira, J. Chanussot, M. D. Mura, W. R. Schwartz, and J. A. d. Santos, "An introduction to deep morphological networks," *arXiv preprint arXiv:1906.01751*, 2019.

[261] J. Masci, J. Angulo, and J. Schmidhuber, "A learning framework for morphological operators using counter–harmonic mean," in *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing.* Springer, 2013, Conference Proceedings, pp. 329–340.

[262] X. Shen and W. Bao, "Hyperspectral endmember extraction using spatially weighted simplex strategy," *Remote Sensing*, vol. 11, no. 18, p. 2147, 2019.

[263] G. Franchi, A. Fehri, and A. Yao, "Deep morphological networks," *Pattern Recognition*, p. 107246, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320320300522

[264] Y. Zhang, S. Blusseau, S. Velasco-Forero, I. Bloch, and J. Angulo, "Max-plus operators applied to filter selection and model pruning in neural networks," in *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing.* Springer, 2019, Conference Proceedings, pp. 310–322, https://github.com/yunxiangzhang/Maxplus-Operators.

[265] M. A. Islam, B. Murray, A. Buck, D. T. Anderson, G. Scott, M. Popescu, and J. Keller, "Deep morphological hit-or-miss transform neural network," *arXiv preprint arXiv:1912.02259*, 2019.

[266] R. Mondal, P. Purkait, S. Santra, and B. Chanda, "Morphological networks for image de-raining," in *International Conference on Discrete Geometry for Computer Imagery.* Springer, 2019, Conference Proceedings, pp. 262–275, https://github.com/ranjanZ/2D-Morphological-Network.

[267] M. Sangalli and M. E. Valle, "Approaches to multivalued mathematical morphology based on uncertain reduced orderings," in *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*. Springer, 2019, Conference Proceedings, pp. 228–240.

[268] B. Burgeth and A. Kleefeld, "An approach to color-morphology based on einstein addition and loewner order," *Pattern Recognition Letters*, vol. 47, no. Supplement C, pp. 29–39, 2014.

[269] G.-S. Xia, X. Bai, J. Ding, Z. Zhu, S. Belongie, J. Luo, M. Datcu, M. Pelillo, and L. Zhang, "Dota: A large-scale dataset for object detection in aerial images," in *Proc. CVPR*, 2018, Conference Proceedings.

[270] "Endmember induction algorithms (eias) toolbox." [Online]. Available: http://www.ehu.es/computationalintelligence/index.php/Endmember_Induction_Algorithms

[271] R. F. Kokaly, R. N. Clark, G. A. Swayze, K. E. Livo, T. M. Hoefen, N. C. Pearson, R. A. Wise, W. M. Benzel, H. A. Lowers, and R. L. Driscoll, "Usgs spectral library version 7," US Geological Survey, Report 2327-638X, 2017.

[272] "Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios Rec. ITU-R BT.601-7," International Telecommunications Union, Geneva, Switzerland, March 2011.

[273] Y. Inoue, "Satellite-and drone-based remote sensing of crops and soils for smart farming–a review," *Soil Science and Plant Nutrition*, vol. 66, no. 6, pp. 798–810, 2020.

[274] R. K. Goel, C. S. Yadav, S. Vishnoi, and R. Rastogi, "Smart agriculture–urgent need of the day in developing countries," *Sustainable Computing: Informatics and Systems*, vol. 30, p. 100512, 2021.

[275] J. A. Van Aardt, D. McKeown, J. Faulring, N. Raqueño, M. Casterline, C. Renschler, R. Eguchi, D. Messinger, R. Krzaczek, and S. Cavillia, "Geospatial disaster response during the haiti earthquake: A case study spanning airborne deployment, data collection, transfer, processing, and dissemination," *Photogrammetric engineering and remote sensing*, vol. 77, no. 9, pp. 943–952, 2011.

[276] C. Kwan, B. Chou, J. Yang, and T. Tran, "Compressive object tracking and classification using deep learning for infrared videos," in *Pattern Recognition and Tracking XXX*, vol. 10995. International Society for Optics and Photonics, 2019, Conference Proceedings, p. 1099506.

[277] J. Theiler, M. Kucer, and A. Ziemann, "Experiments in anomalous change detection with the viareggio 2013 trial dataset," in *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imagery XXVI*, vol. 11392. International Society for Optics and Photonics, 2020, Conference Proceedings, p. 1139211.

[278] C.-I. Chang, *Hyperspectral target detection*. Springer, 2016, pp. 131–172.

[279] Q. Wang, Q. Lin, M. Li, and Q. Tian, "A new target detection algorithm: spectra sort encoding," *International Journal of Remote Sensing*, vol. 30, no. 9, pp. 2297–2307, 2009. [Online]. Available: https://doi.org/10.1080/01431160802549351

[280] C. Chang, H. Cao, S. Chen, X. Shang, C. Yu, and M. Song, "Orthogonal subspace projection-based go-decomposition approach to finding low-rank and sparsity matrices for hyperspectral anomaly detection," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–27, 2020.

[281] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[282] E. Vargas, S. Pinilla, and H. Arguello, "A fast endmember estimation algorithm from compressive measurements," 2018.

[283] H. Vargas, Y. Fonseca, and H. Arguello, "Object detection on compressive measurements using correlation filters and sparse representation," 2018.

[284] T. Gelvez and H. Arguello, "Spectral image fusion from compressive projections using total-variation and low-rank regularizations," 2018.

[285] H. Rueda, H. Arguello, and G. R. Arce, "High-dimensional optimization of color coded apertures for compressive spectral cameras," in *Suropean Signal Processing Conference (EUSIPCO)*, 2017, Conference Proceedings.

[286] J. M. Ramirez, J. I. M. Torre, and H. Arguello, "Feature fusion via dual-resolution compressive measurement matrix analysis for spectral image classification," *Signal Processing: Image Communication*, vol. 90, p. 116014, 2021.

[287] M. Marquez, H. Rueda-Chacon, and H. Arguello, "Compressive spectral imaging via virtual side information," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 114–123, 2021.

[288] D. Kittle, K. Choi, A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," *Applied optics*, vol. 49, no. 36, pp. 6824–6833, 2010.

[289] H. Zhang, X. Ma, X. Zhao, and G. R. Arce, "Compressive hyperspectral image classification using a 3d coded convolutional neural network," *Opt. Express*, vol. 29, no. 21, pp. 32 875–32 891, Oct 2021. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-29-21-32875

[290] D. Kuzin, O. Isupova, and L. Mihaylova, "Compressive sensing approaches for autonomous object detection in video sequences," in *2015 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*. IEEE, 2015, Conference Proceedings, pp. 1–6.

[291] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2241–2253, 2010.

[292] C. S. McCamy, H. Marcus, J. G. Davidson *et al.*, "A color-rendition chart," *J. App. Photog. Eng*, vol. 2, no. 3, pp. 95–99, 1976.

[293] RIT, "Munsell color science lab educational resources." [Online]. Available: https://www.rit.edu/science/munsell-color-science-lab-educational-resources

Appendix A

Datasets

A.1 Dataset of Object deTection in Aerial images (DOTA)

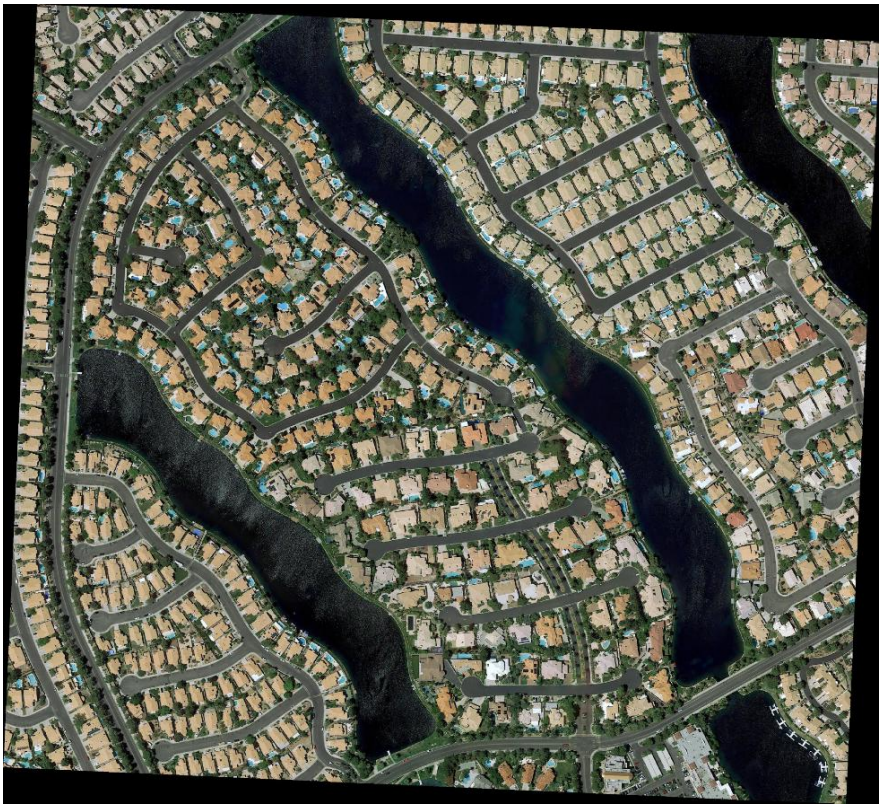


Figure A.1.1: DOTA Image P0059 (*Image 1*) – 36.210048, -115.271147

Figure A.1.2: DOTA Image P0060 (*Image 2*) – 36.235162, -115.265200

Figure A.1.3: DOTA Image P0061 (*Image 3*) – 36.108970, -115.240463

Figure A.1.4: DOTA Image P0062 (*Image 4*) – 36.111525, -115.235634

Figure A.1.5: DOTA Image P0063 (*Image 5*) – 36.119661, -115.212777

Figure A.1.6: DOTA Image P0064 (*Image 6*) – 36.154106, -115.182226

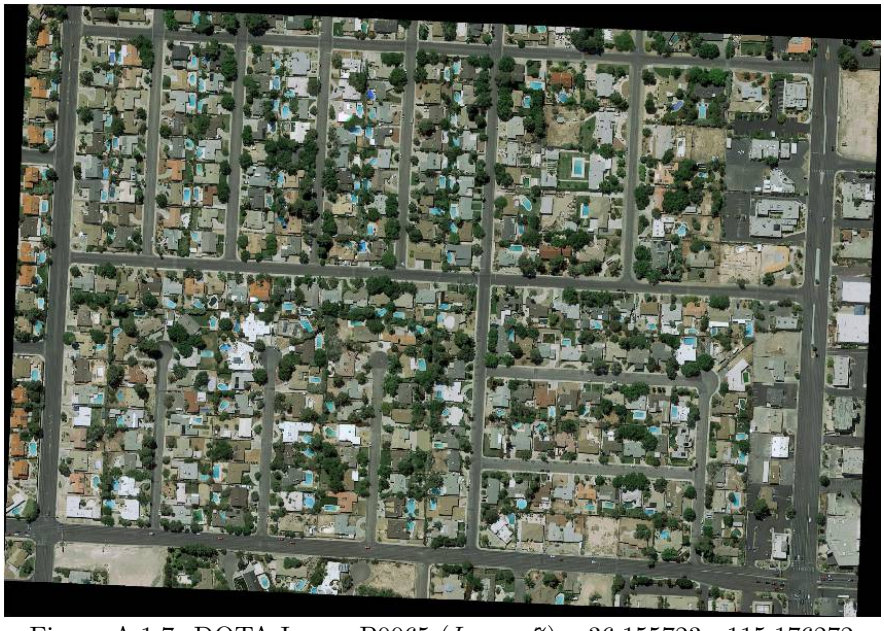


Figure A.1.7: DOTA Image P0065 (*Image 7*) – 36.155723, -115.176272

A.2 CAVE GAP Camera Multispectral Image Database

(a) (b)

(c) (d)

(e) (f)

(g) (h)

Figure A.2.1: Subset of the CAVE multispectral dataset used to test simulated CASSI TD architectures. **a)** Balloons, **b)** beads **c)** clay, **d)** feathers, **e)** glass tiles, **f)** pompoms, **g)** sponges, and **h)** superballs.

A.3 Hyperspectral Image Datasets

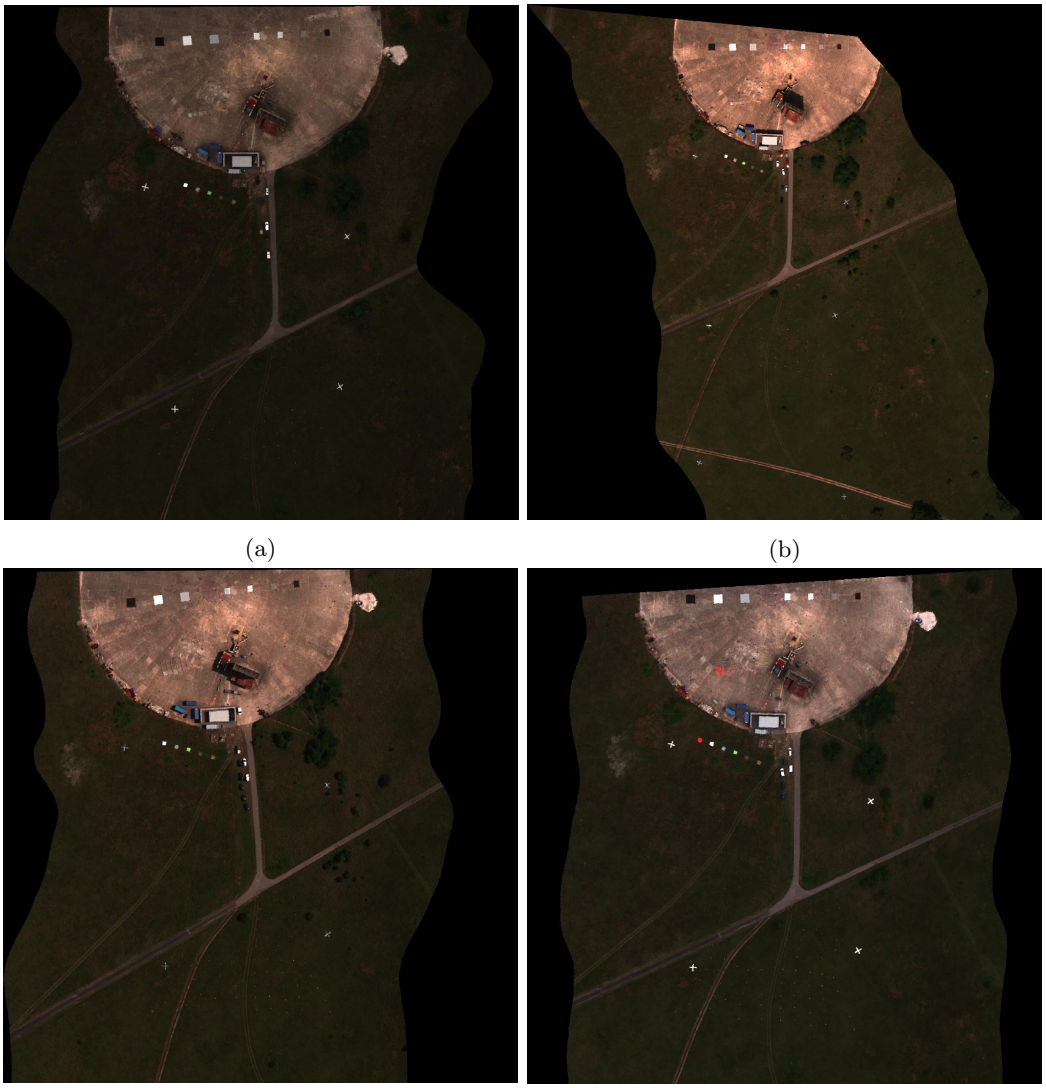
A.3.1 "Operation 7"

(a) (b) (c)

Figure A.3.1: Images from the OP7 Hyperspectral Image Dataset provided by BAE Systems[1].

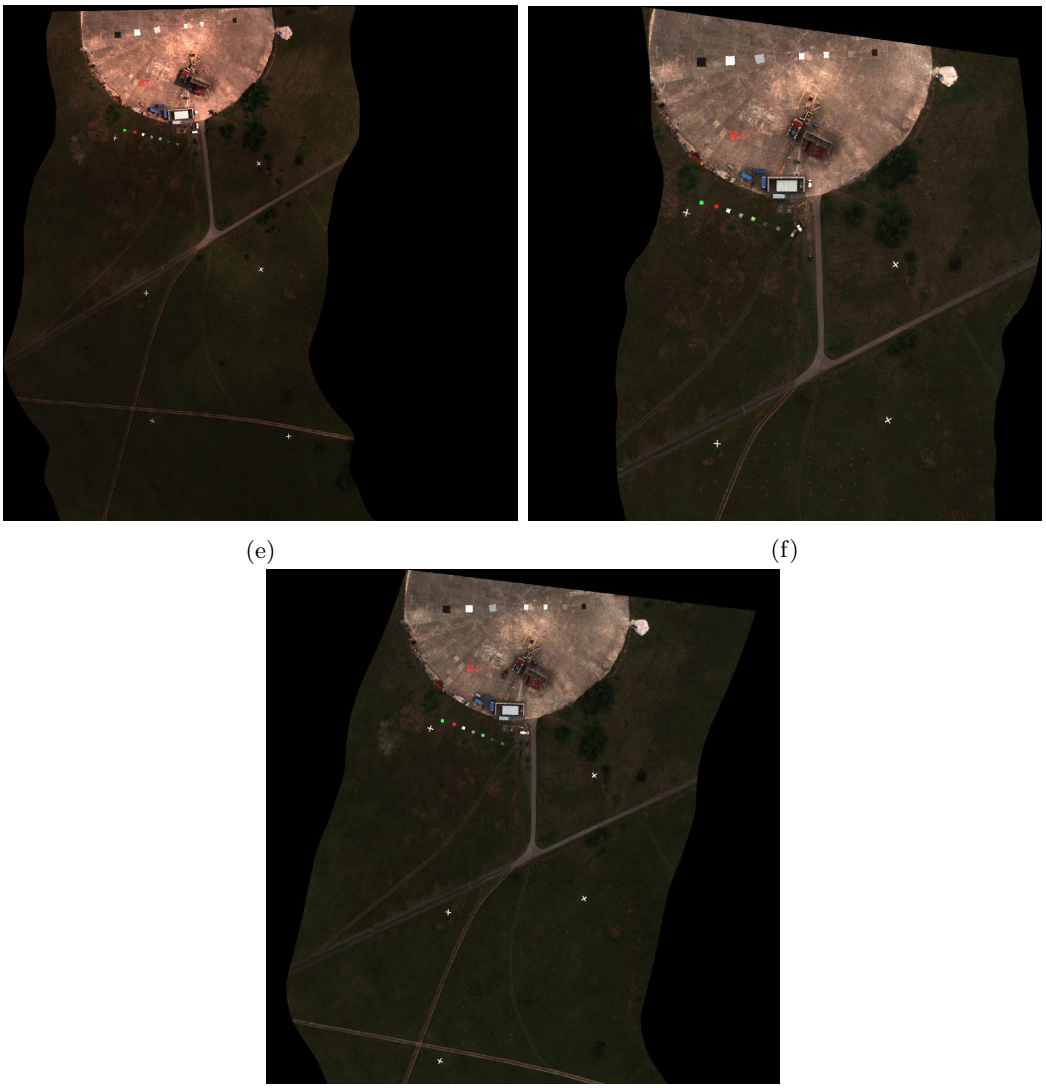
[1]Copyright ©2025 BAE Systems. All Rights Reserved. BAE SYSTEMS is a registered Trademark of BAE Systems plc.

A.3.2 UDRC Selene Trial



(a) (b)

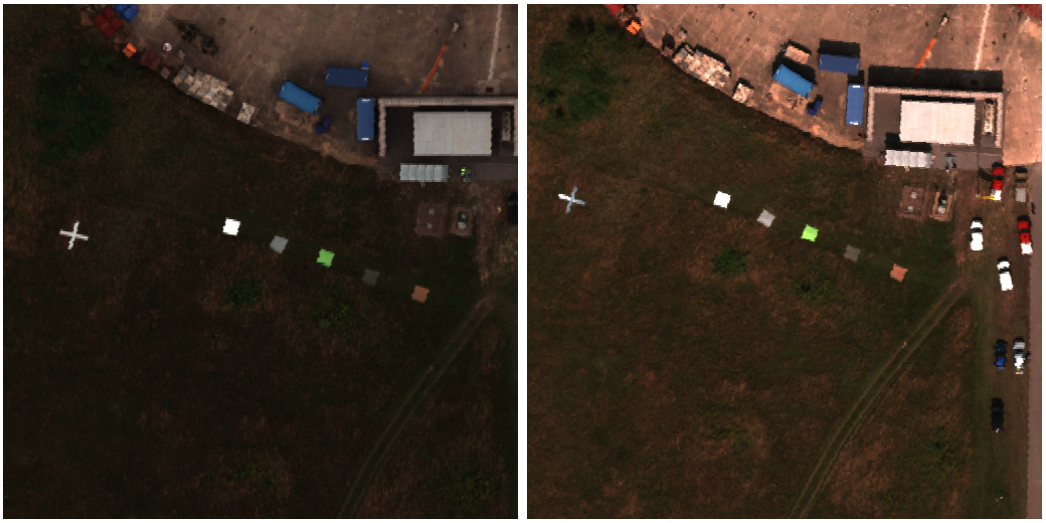
(c) (d)



(e) (f)

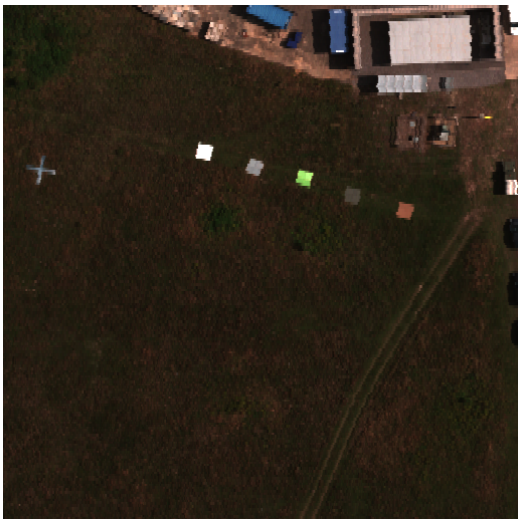
(g)

Figure A.3.2: Subsection of Hyperspectral Images from the UDRC Selene Trial provided by DSTL. **a)** IM140804, **b)** IM140806, **c)** IM140807, **d)** IM140808, **e)** IM140812, **f)** IM140813, **g)** IM140815.



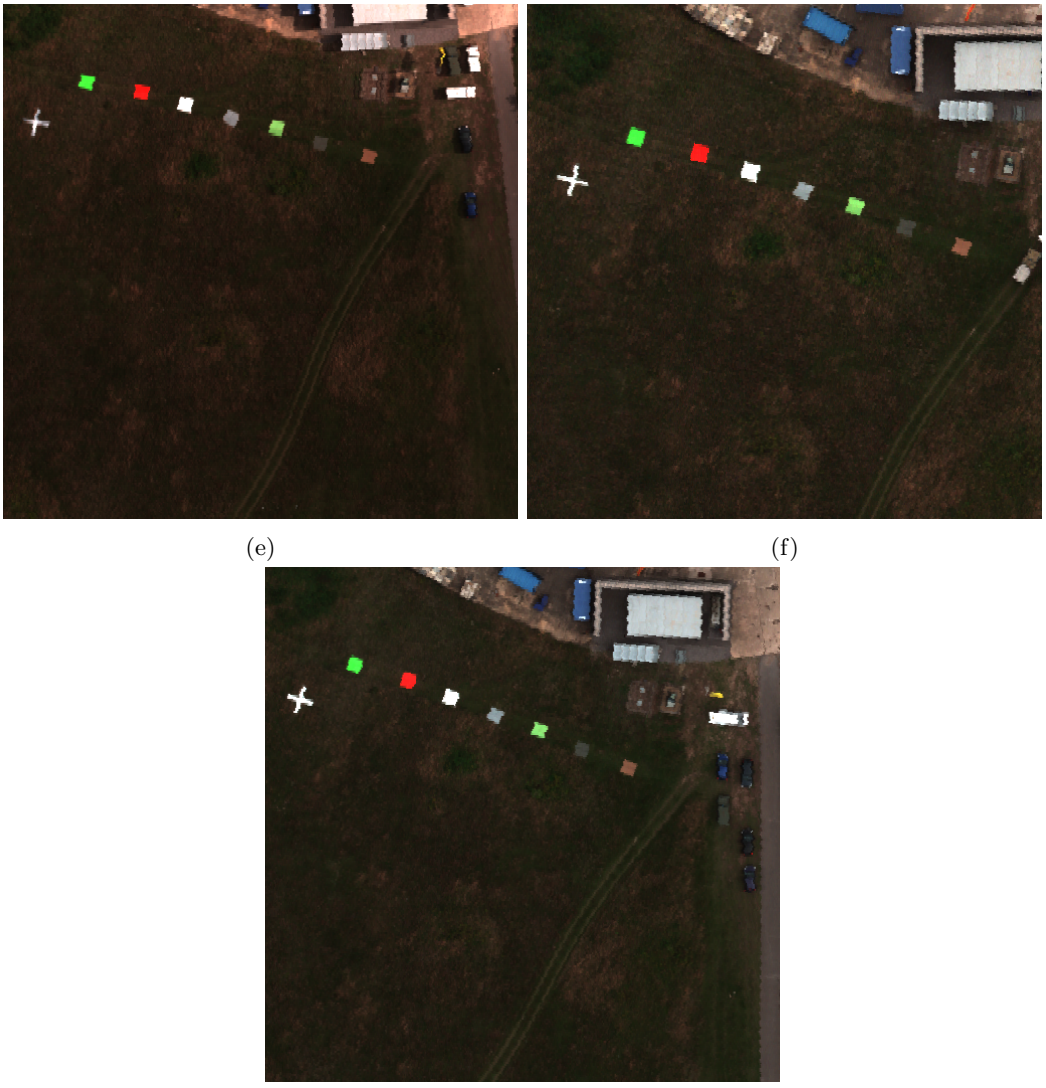
(a)

(b)



(c)

(d)



(e) (f)

(g)

Figure A.3.3: Cropped subsection of Hyperspectral Images from the UDRC Selene Trial provided by DSTL used with the JSSDR-TD and CASSI-TD architectures. **a)** IM140804, **b)** IM140806, **c)** IM140807, **d)** IM140808, **e)** IM140812, **f)** IM140813, **g)** IM140815.

Appendix B

Supplementary Results

B.1 Additional MDPOHMT results on synthetic images

In order to investigate the effects of the background colour on the performance of the MDPOHMT in colour images multiple block colour backgrounds were used similar to the ones from Figure 4.4.

Table B.1: Red, Green, and Blue values for the synthetic image background colours

BG Colour	[R, G, B]
Orange	[255, 165, 55]
Magenta	[255, 55, 165]
Green	[165, 255, 55]
Cyan	[55, 255, 165]
Purple	[165, 55, 255]
Blue	[55, 165, 255]

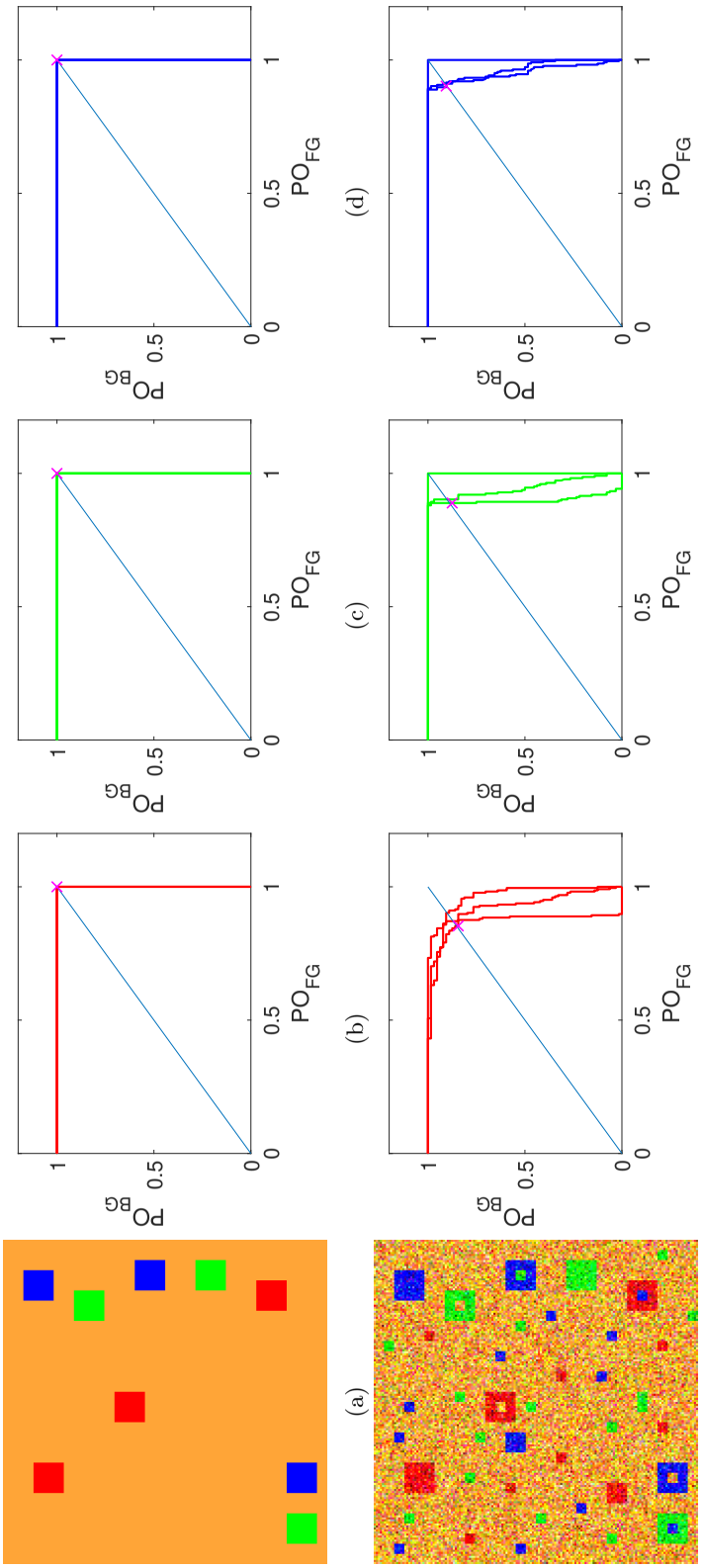


Figure B.1.1: Target object PO plots obtained by applying the MDPOHMT with an orange background

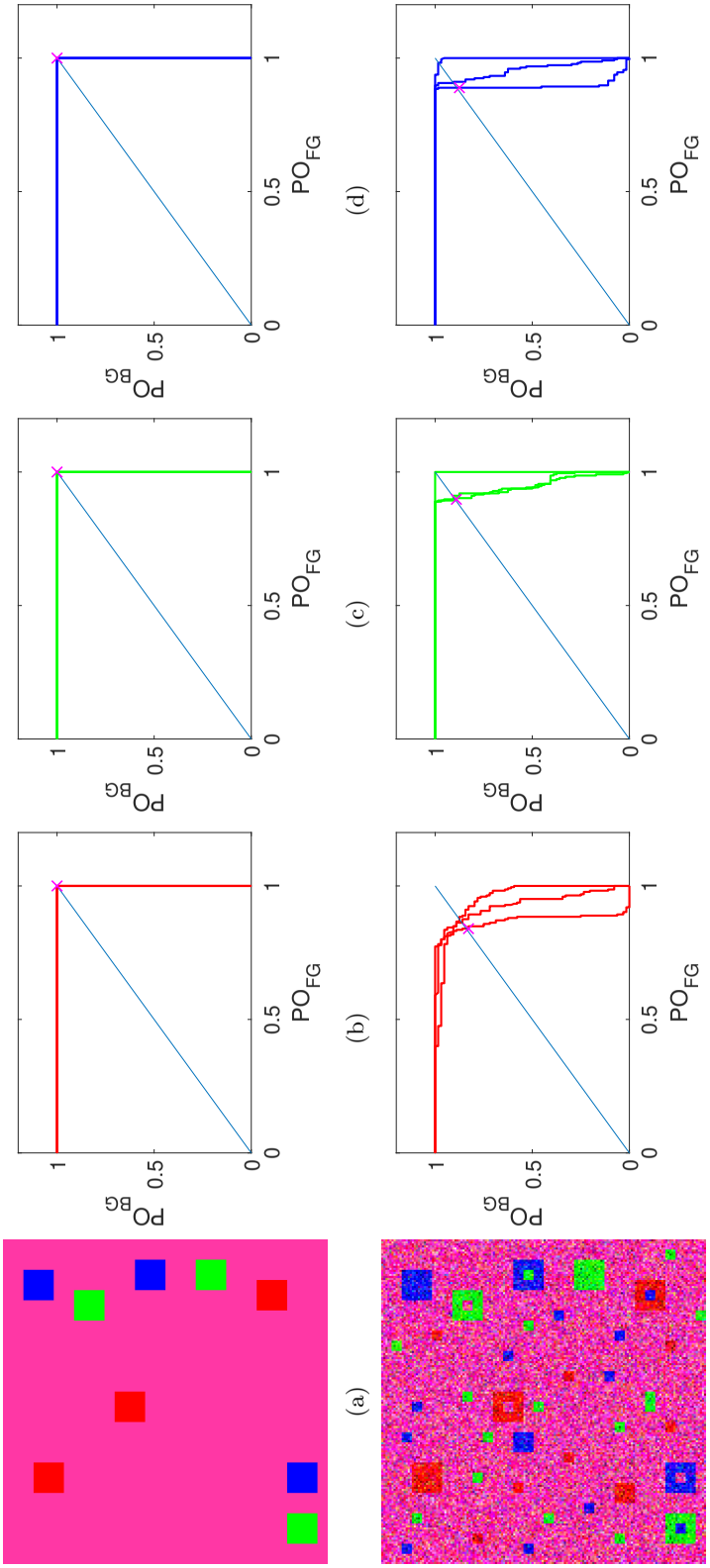


Figure B.1.2: Target object PO plots obtained by applying the MDPOHMT with a magenta background

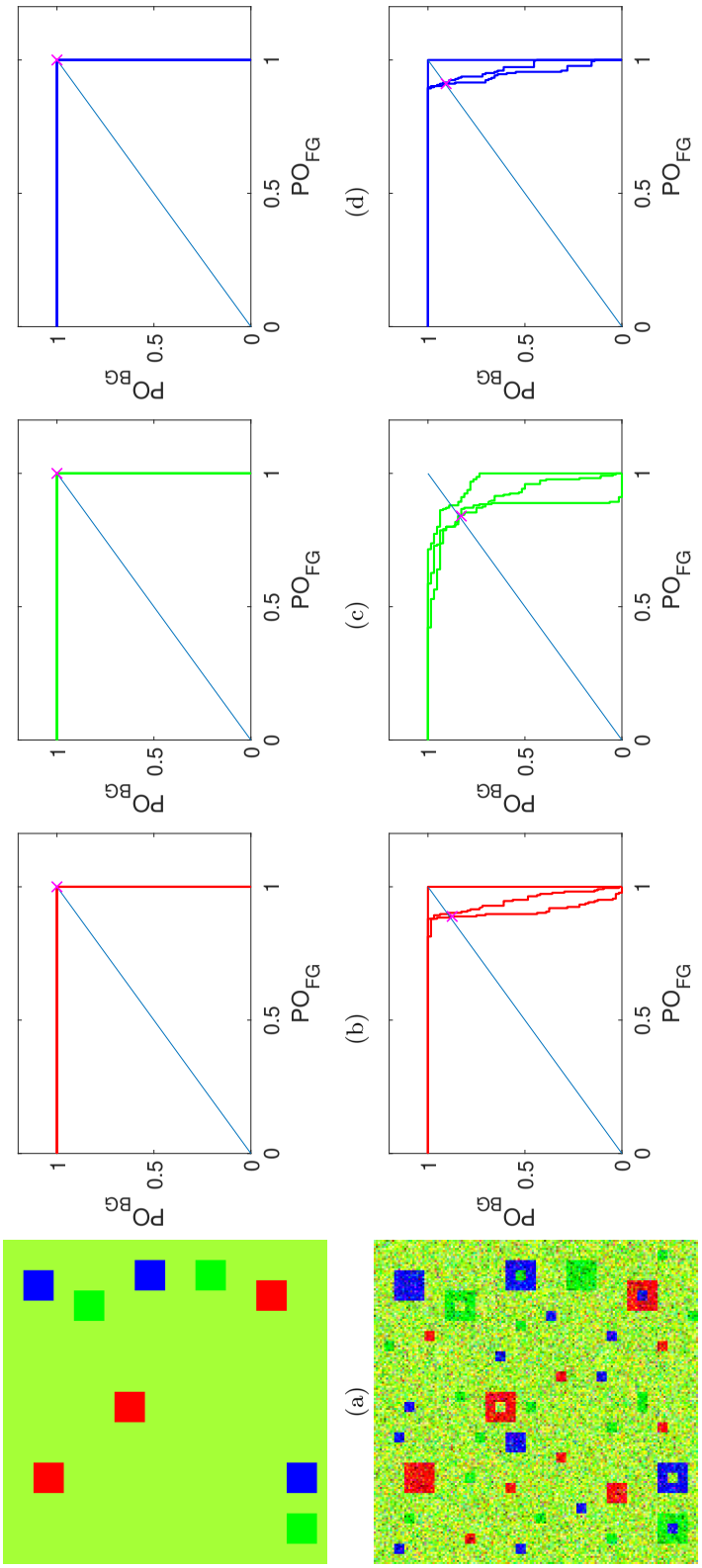


Figure B.1.3: Target object PO plots obtained by applying the MDPOHMT with a green background

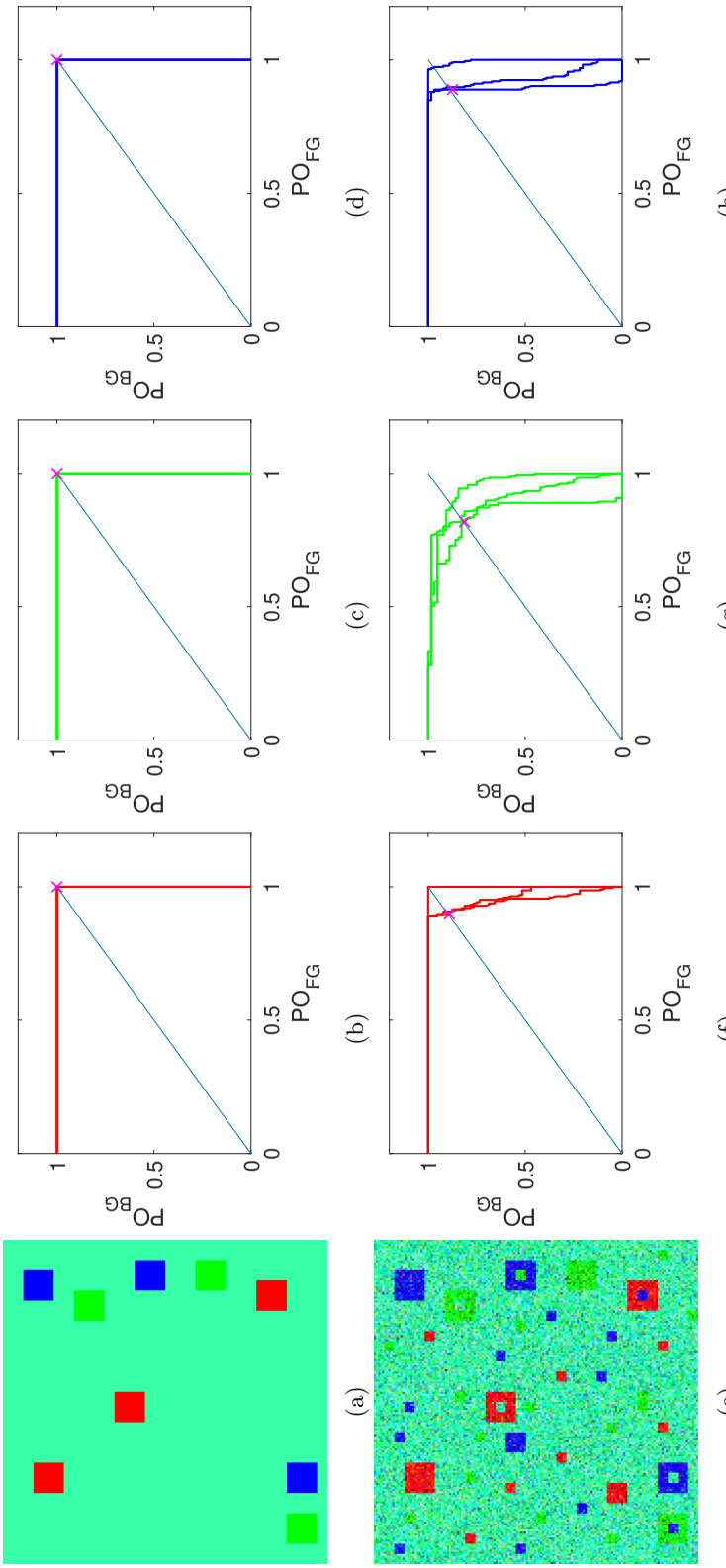


Figure B.1.4: Target object PO plots obtained by applying the MDPOHMT with a cyan background

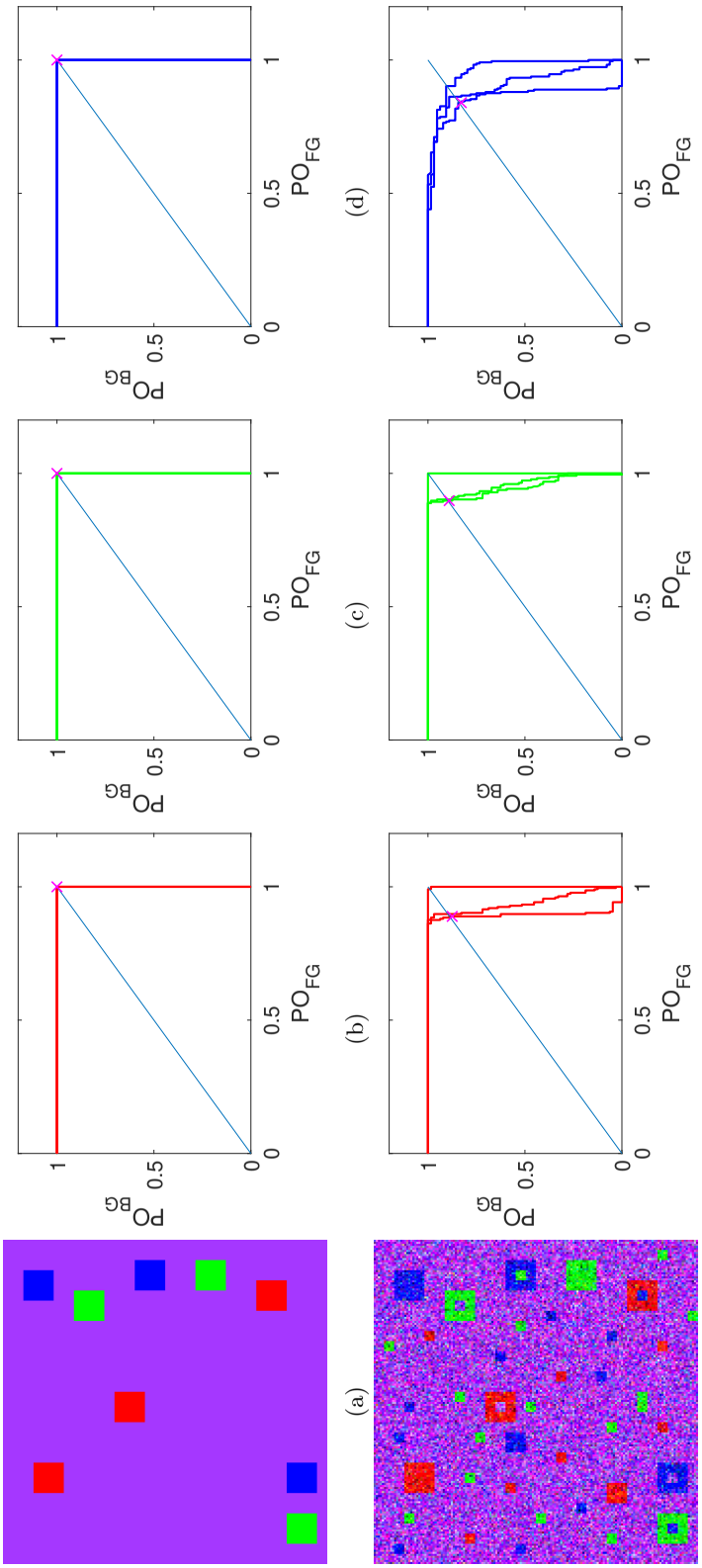


Figure B.1.5: Target object PO plots obtained by applying the MDPOHMT with a purple background

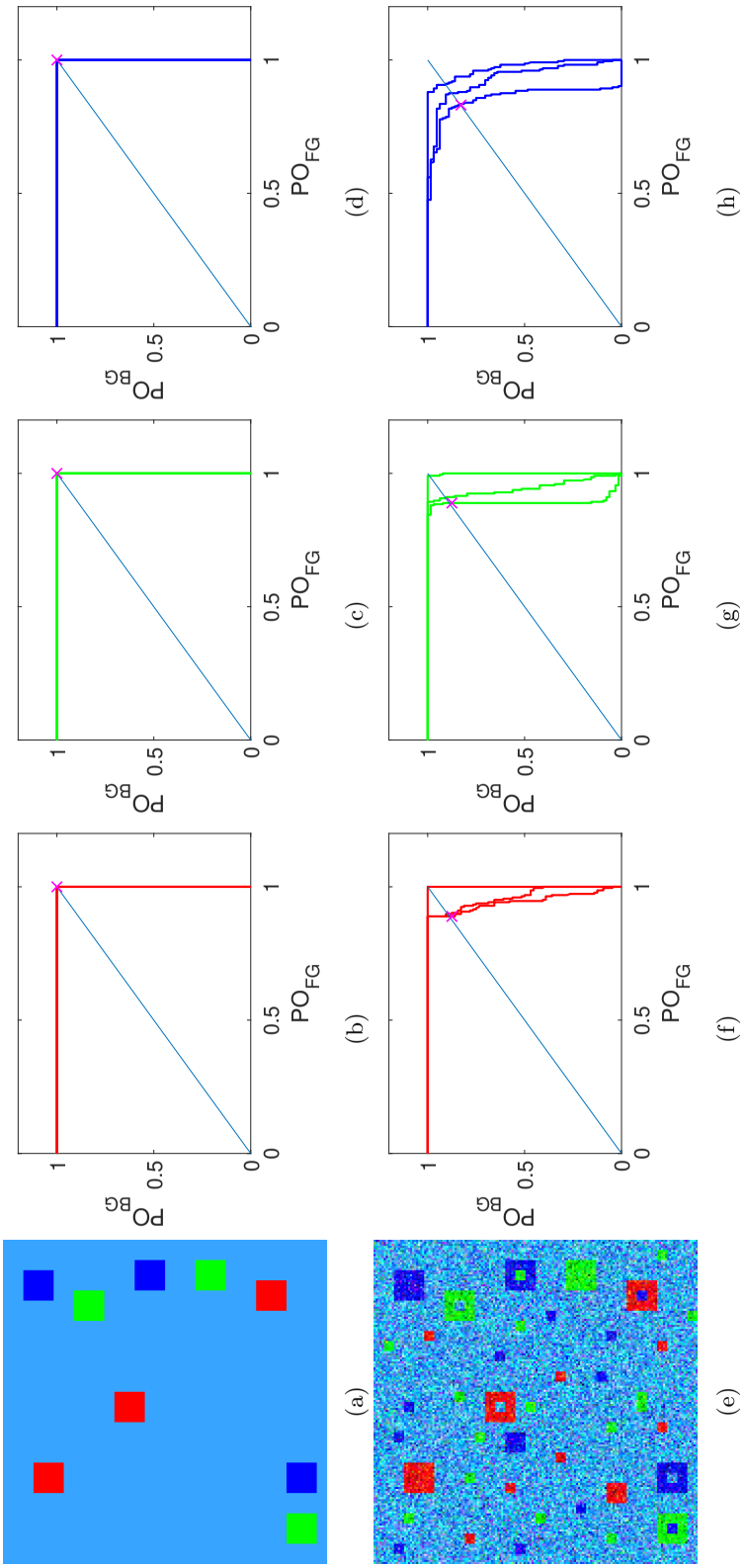


Figure B.1.6: Target object PO plots obtained by applying the MDPOHMT with a blue background

B.2 Additional UHMT Results

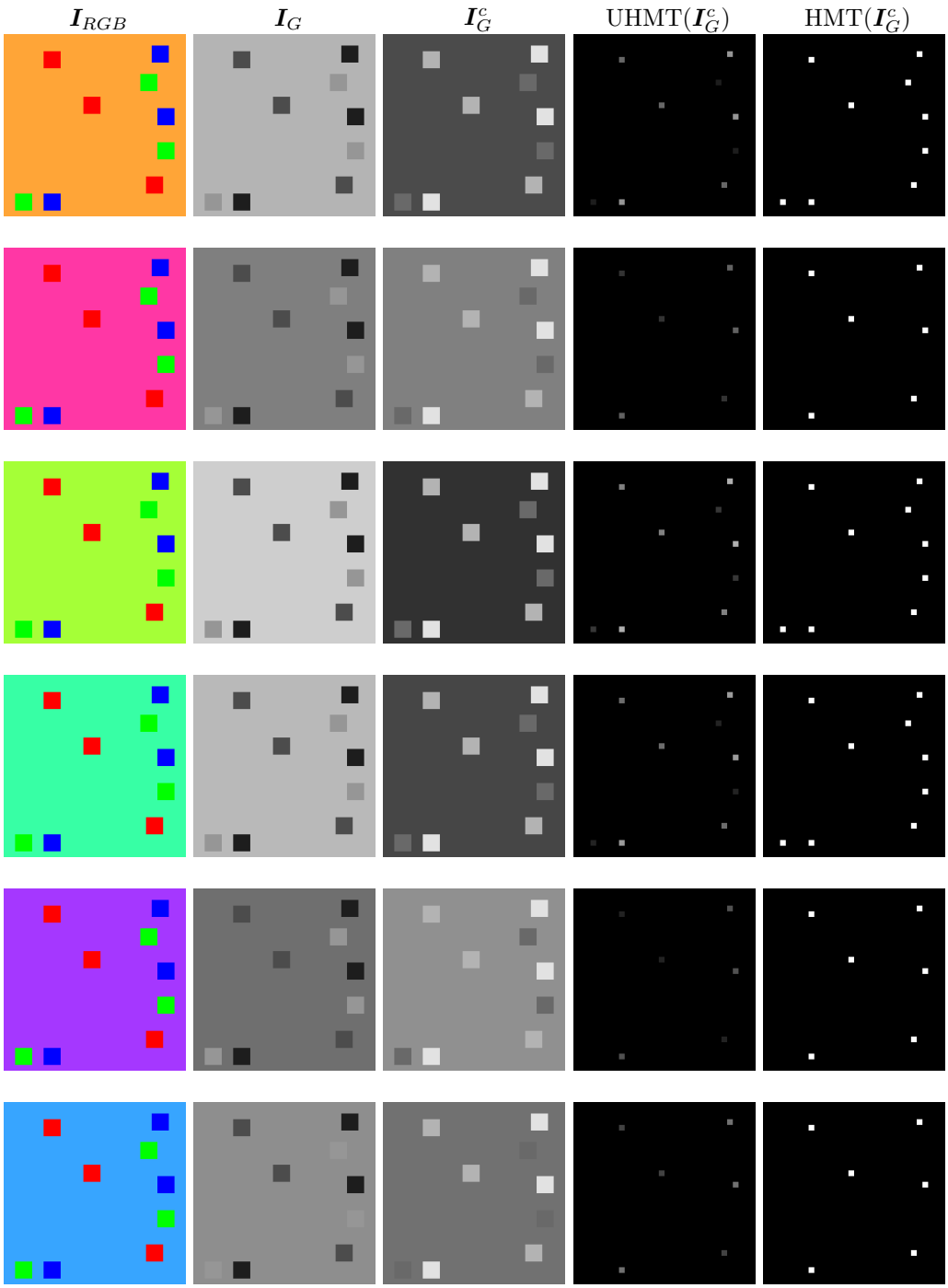
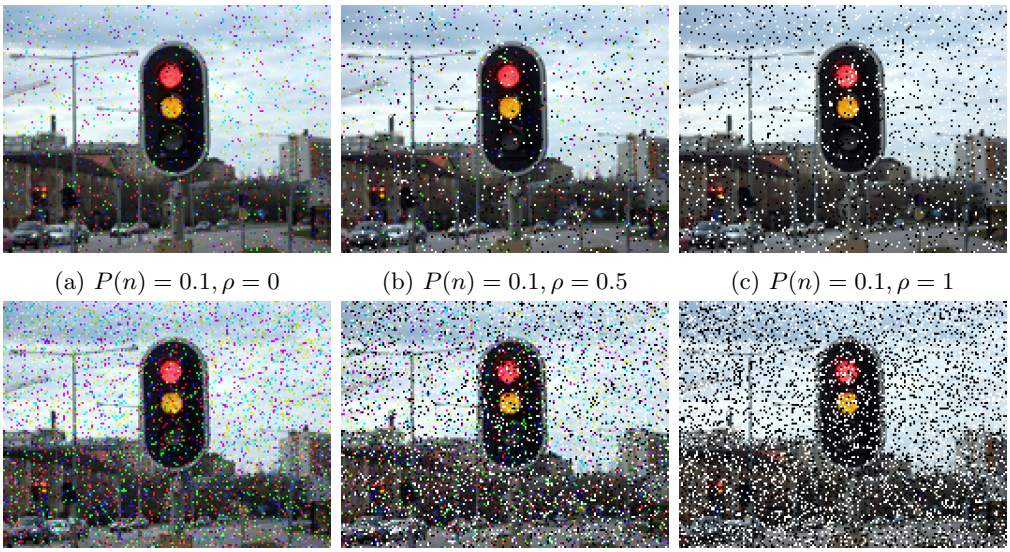


Figure B.2.1: Additional Greyscale UHMT and HMT Results on Synthetic Colour Images

B.3 MDPOHMT Results on Noise Compromised Images

B.3.1 Images

B.3.1.1 Bipolar Impulsive Noise

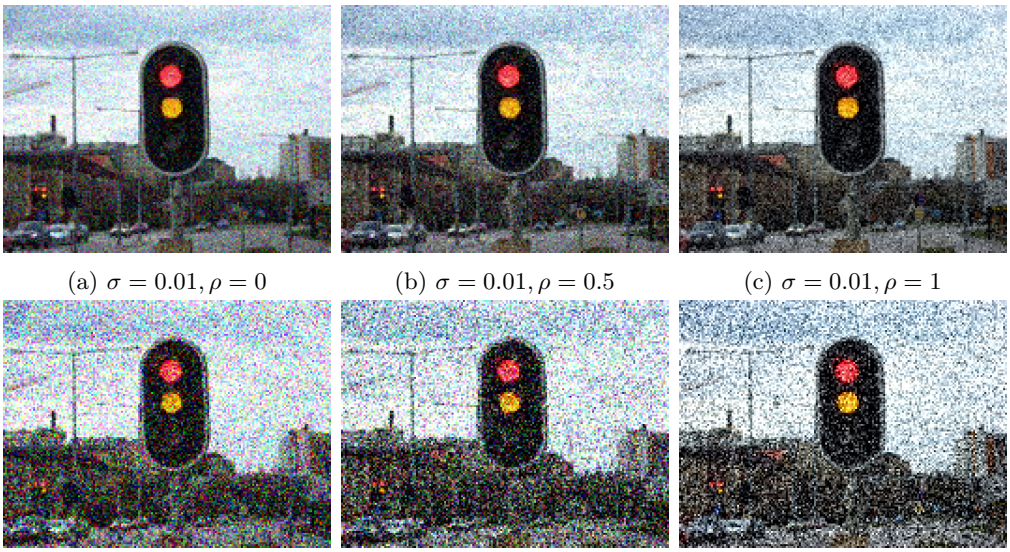


(a) $P(n) = 0.1, \rho = 0$ (b) $P(n) = 0.1, \rho = 0.5$ (c) $P(n) = 0.1, \rho = 1$

(d) $P(n) = 0.3, \rho = 0$ (e) $P(n) = 0.3, \rho = 0.5$ (f) $P(n) = 0.3, \rho = 1$

Figure B.3.1: Traffic light images with added Bipolar Impulsive Noise.

B.3.1.2 Gaussian Noise

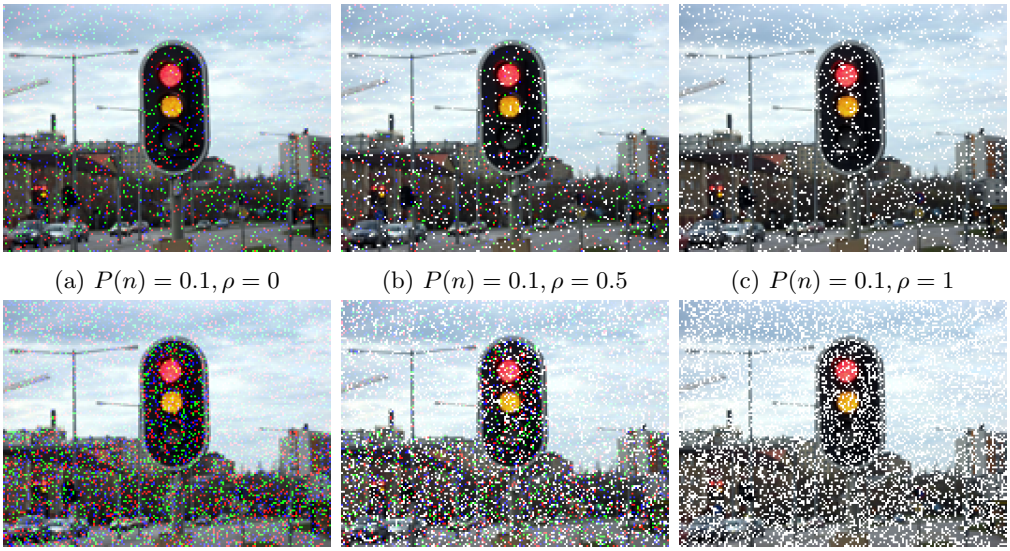


(a) $\sigma = 0.01, \rho = 0$ (b) $\sigma = 0.01, \rho = 0.5$ (c) $\sigma = 0.01, \rho = 1$

(d) $\sigma = 0.05, \rho = 0$ (e) $\sigma = 0.05, \rho = 0.5$ (f) $\sigma = 0.05, \rho = 1$

Figure B.3.2: Traffic light images with added Gaussian Noise.

B.3.1.3 Impulsive Noise



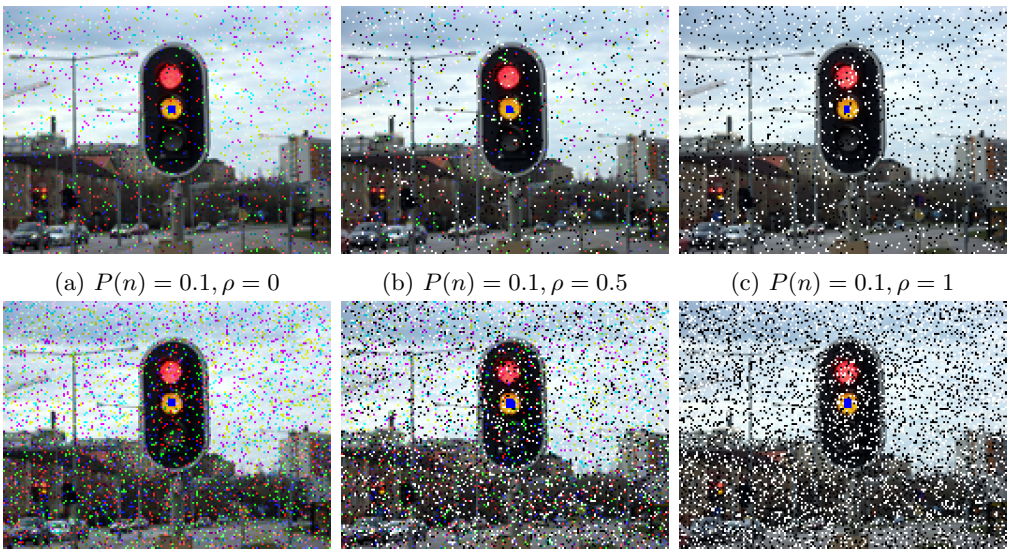
(a) $P(n) = 0.1, \rho = 0$ (b) $P(n) = 0.1, \rho = 0.5$ (c) $P(n) = 0.1, \rho = 1$

(d) $P(n) = 0.3, \rho = 0$ (e) $P(n) = 0.3, \rho = 0.5$ (f) $P(n) = 0.3, \rho = 1$

Figure B.3.3: Traffic light images with added Impulsive Noise.

B.3.2 MDPOHMT

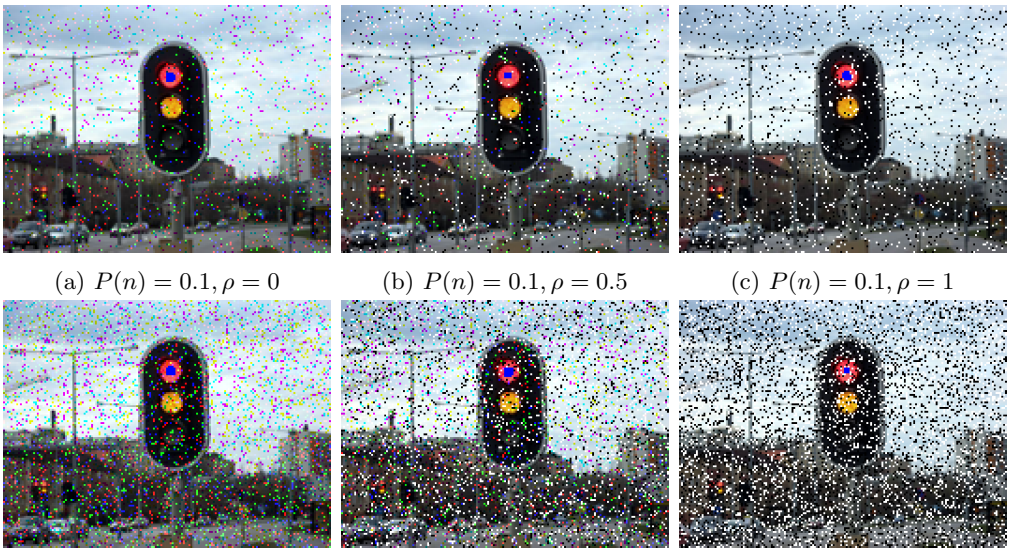
B.3.2.1 Bipolar Impulsive Noise



(a) $P(n) = 0.1, \rho = 0$ (b) $P(n) = 0.1, \rho = 0.5$ (c) $P(n) = 0.1, \rho = 1$

(d) $P(n) = 0.3, \rho = 0$ (e) $P(n) = 0.3, \rho = 0.5$ (f) $P(n) = 0.3, \rho = 1$

Figure B.3.4: MDPOHMT results in detecting the amber lights in the Bipolar Impulsive Noise compromised images from Figure B.3.1.

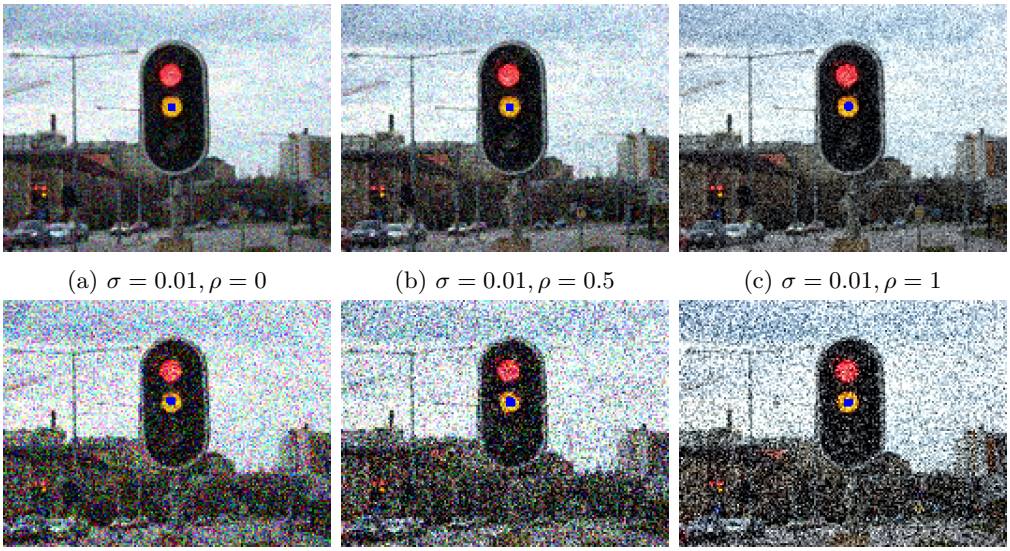


(a) $P(n) = 0.1, \rho = 0$ (b) $P(n) = 0.1, \rho = 0.5$ (c) $P(n) = 0.1, \rho = 1$

(d) $P(n) = 0.3, \rho = 0$ (e) $P(n) = 0.3, \rho = 0.5$ (f) $P(n) = 0.3, \rho = 1$

Figure B.3.5: MDPOHMT results in detecting the red lights in the Bipolar Impulsive Noise compromised images from Figure B.3.1.

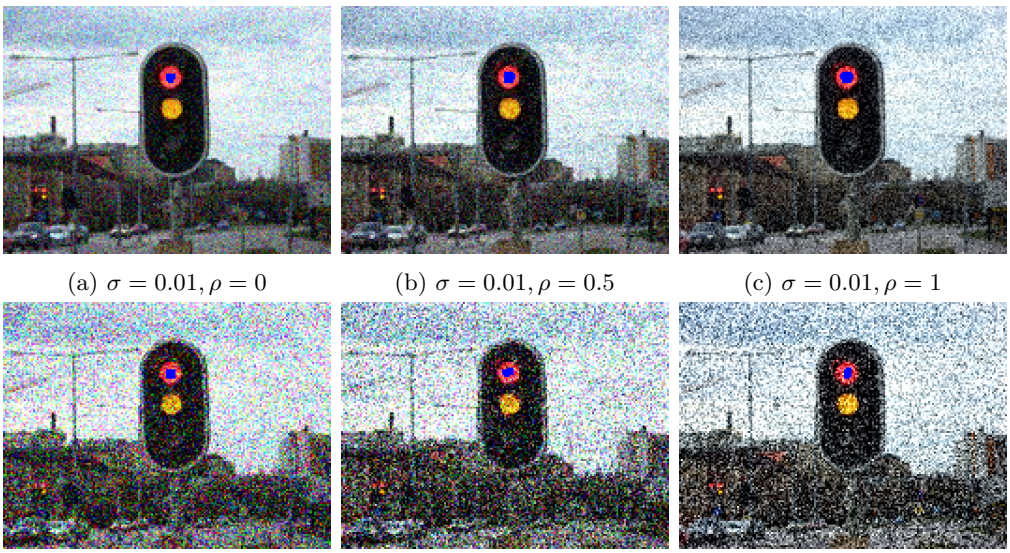
B.3.2.2 Gaussian Noise



(a) $\sigma = 0.01, \rho = 0$ (b) $\sigma = 0.01, \rho = 0.5$ (c) $\sigma = 0.01, \rho = 1$

(d) $\sigma = 0.05, \rho = 0$ (e) $\sigma = 0.05, \rho = 0.5$ (f) $\sigma = 0.05, \rho = 1$

captionMDPOHMT results in detecting the amber lights in the Gaussian Noise compromised images from Figure B.3.2.

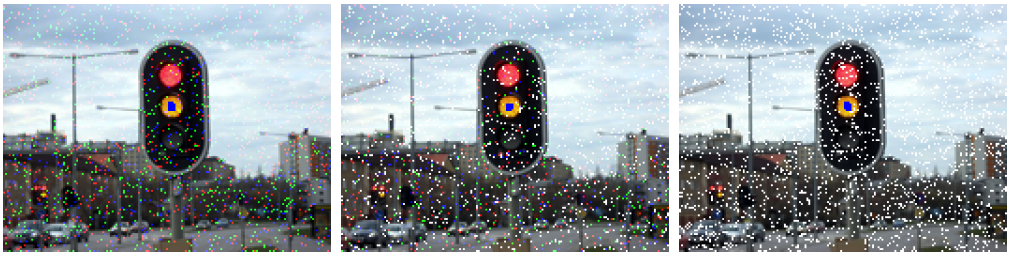


(a) $\sigma = 0.01, \rho = 0$ (b) $\sigma = 0.01, \rho = 0.5$ (c) $\sigma = 0.01, \rho = 1$

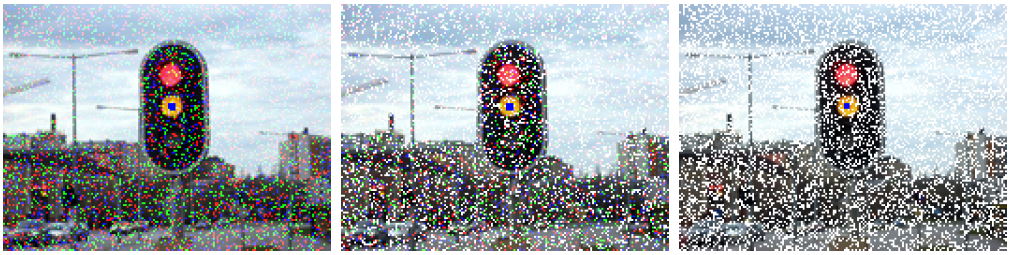
(d) $\sigma = 0.05, \rho = 0$ (e) $\sigma = 0.05, \rho = 0.5$ (f) $\sigma = 0.05, \rho = 1$

Figure B.3.7: MDPOHMT results in detecting the red lights in the Gaussian Noise compromised images from Figure B.3.2.

B.3.2.3 Impulsive Noise

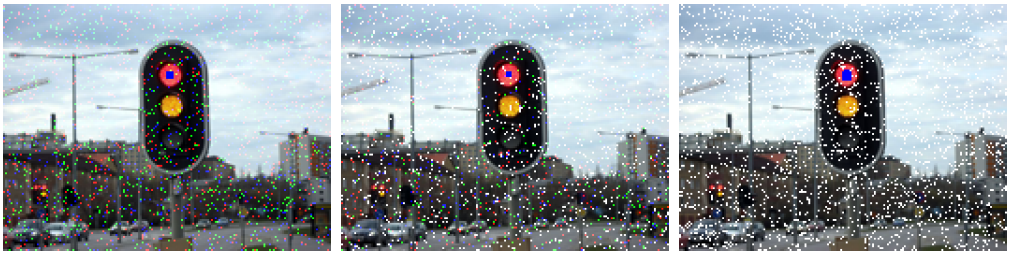


(a) $P(n) = 0.1, \rho = 0$ (b) $P(n) = 0.1, \rho = 0.5$ (c) $P(n) = 0.1, \rho = 1$

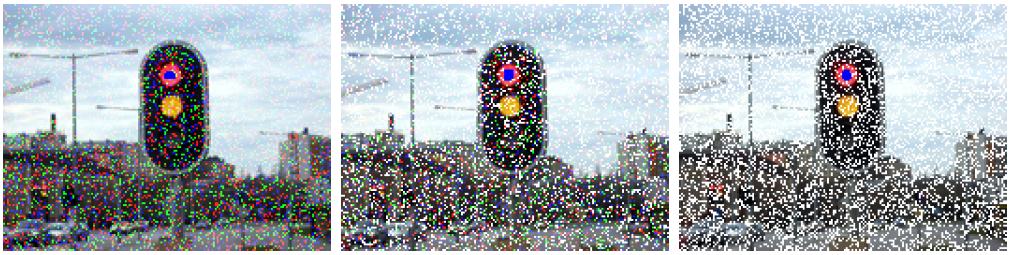


(d) $P(n) = 0.3, \rho = 0$ (e) $P(n) = 0.3, \rho = 0.5$ (f) $P(n) = 0.3, \rho = 1$

Figure B.3.8: MDPOHMT results in detecting the amber lights in the Impulsive Noise compromised images from Figure B.3.3.



(a) $P(n) = 0.1, \rho = 0$ (b) $P(n) = 0.1, \rho = 0.5$ (c) $P(n) = 0.1, \rho = 1$



(d) $P(n) = 0.3, \rho = 0$ (e) $P(n) = 0.3, \rho = 0.5$ (f) $P(n) = 0.3, \rho = 1$

Figure B.3.9: MDPOHMT results in detecting the red lights in the Impulsive Noise compromised images from Figure B.3.3.

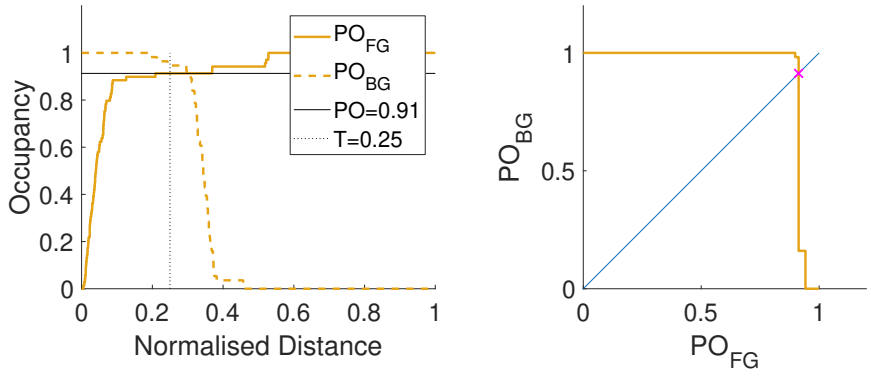
B.3.3 MDPOHMT Parameters

Table B.2: MDPOHMT parameters used in noisy traffic light images.

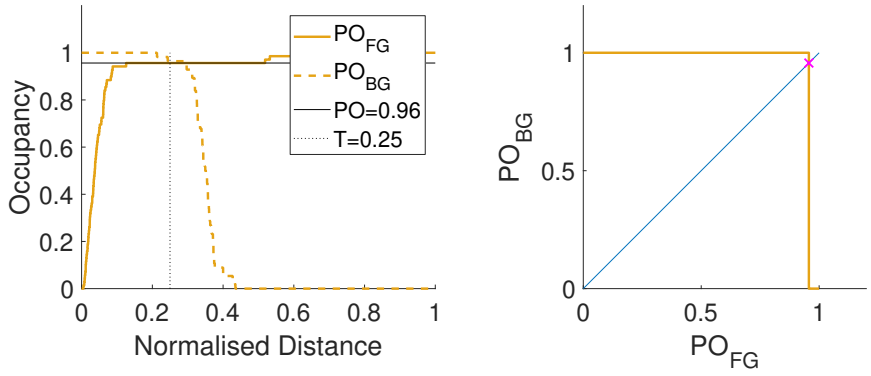
Light	Noise	$P(n)/\sigma$	ρ	Estimated PO	PO_{BG}	PO_{FG}	T
Red	Gaussian	0.01	0	0.99	0.96	0.97	0.25
			0.5	0.91	0.89	0.88	0.25
			1	0.87	0.84	0.86	0.25
		0.05	0	0.75	0.73	0.74	0.3
			0.5	0.75	0.73	0.72	0.25
			1	0.71	0.68	0.70	0.25
	Bipolar Imp.	0.1	0	0.94	0.91	0.93	0.25
			0.5	0.97	0.95	0.96	0.25
			1	0.86	0.82	0.84	0.25
		0.3	0	0.86	0.82	0.84	0.25
			0.5	0.81	0.79	0.80	0.25
			1	0.74	0.71	0.72	0.25
	Impulsive	0.1	0	0.94	0.91	0.93	0.25
			0.5	0.93	0.89	0.91	0.25
			1	0.84	0.82	0.83	0.25
		0.3	0	0.83	0.80	0.81	0.25
			0.5	0.67	0.64	0.65	0.25
			1	0.77	0.75	0.75	0.25
Amber	Gaussian	0.01	0	0.98	0.96	0.96	0.25
			0.5	0.96	0.95	0.94	0.25
			1	0.93	0.89	0.91	0.25
		0.05	0	0.79	0.77	0.77	0.3
			0.5	0.68	0.66	0.65	0.25
			1	0.74	0.71	0.72	0.25
	Bipolar Imp.	0.1	0	0.91	0.89	0.88	0.25
			0.5	0.96	0.93	0.94	0.25
			1	0.90	0.88	0.88	0.25
		0.3	0	0.80	0.77	0.78	0.25
			0.5	0.75	0.73	0.74	0.25
			1	0.71	0.68	0.70	0.25
	Impulsive	0.1	0	0.96	0.93	0.94	0.25
			0.5	0.94	0.91	0.93	0.25
			1	0.90	0.88	0.88	0.25
		0.3	0	0.87	0.84	0.86	0.25
			0.5	0.77	0.75	0.75	0.25
			1	0.72	0.70	0.71	0.25

B.3.4 PO Plots

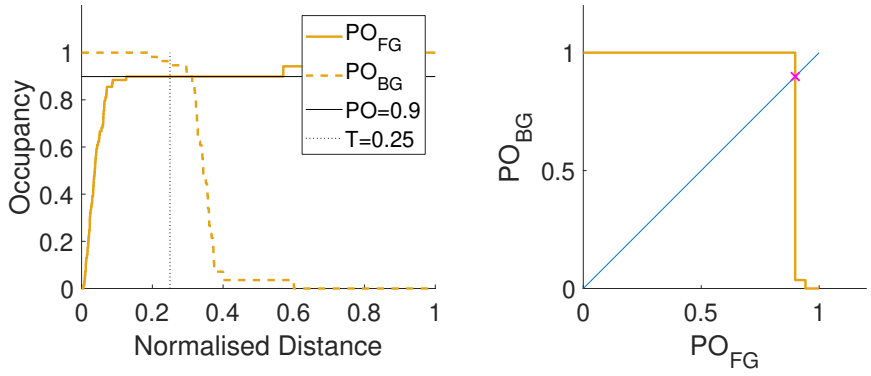
B.3.4.1 Bipolar Impulsive Noise



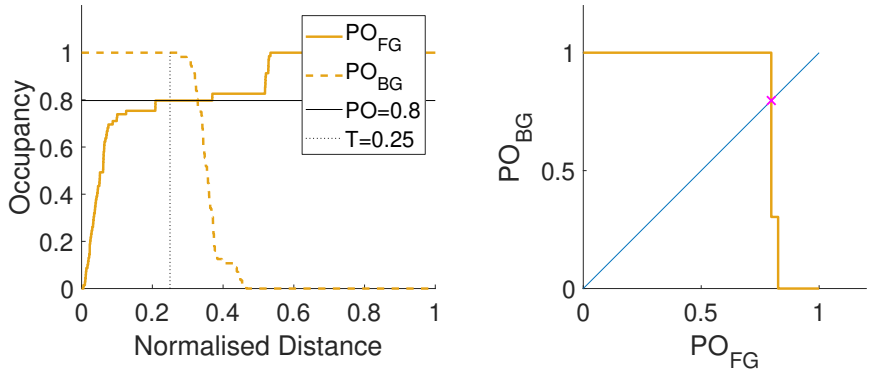
(a) $P(n) = 0.1,\ \rho = 0$



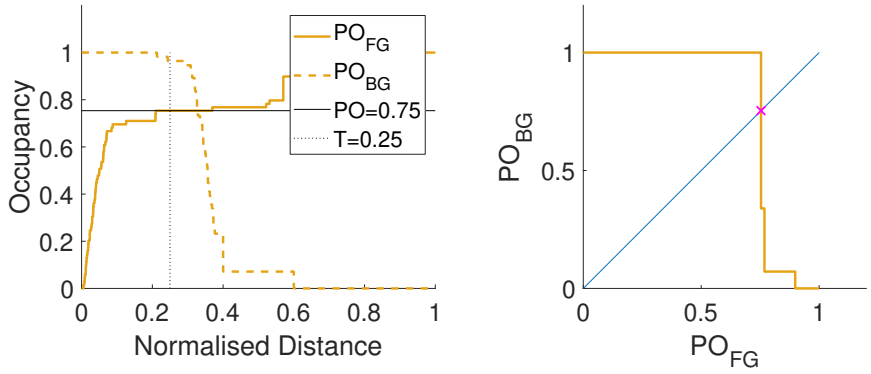
(b) $P(n) = 0.1,\ \rho = 0.5$



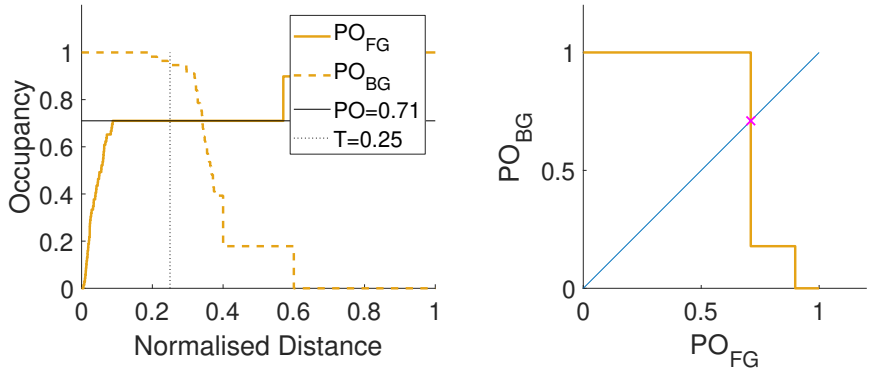
(c) $P(n) = 0.1,\ \rho = 1$



(d) $P(n) = 0.3,\ \rho = 0$

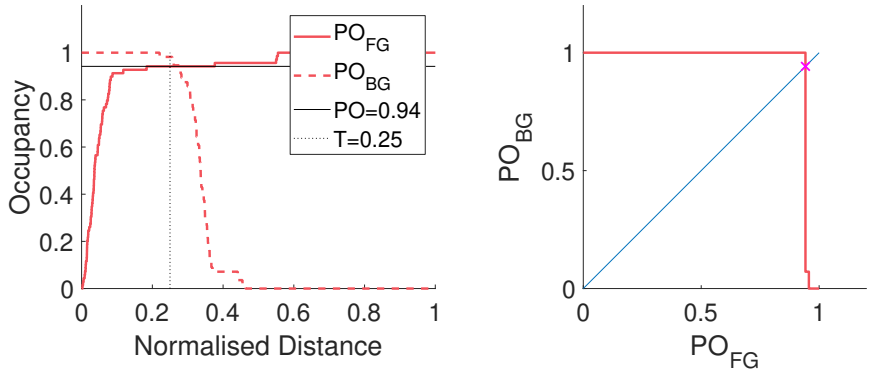


(e) $P(n) = 0.3,\ \rho = 0.5$

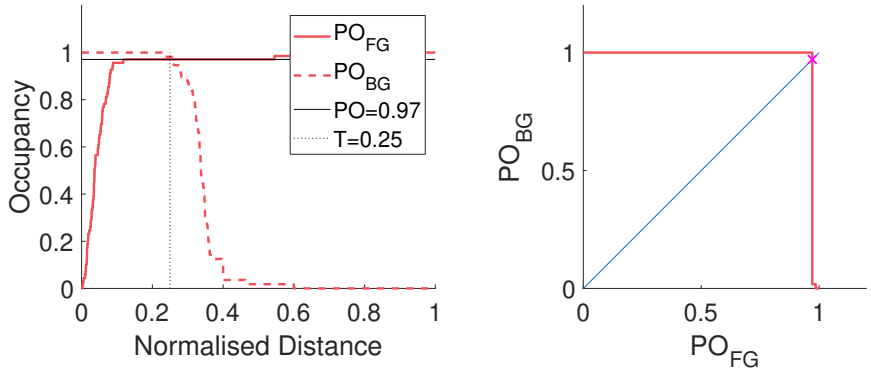


(f) $P(n) = 0.3,\ \rho = 1$

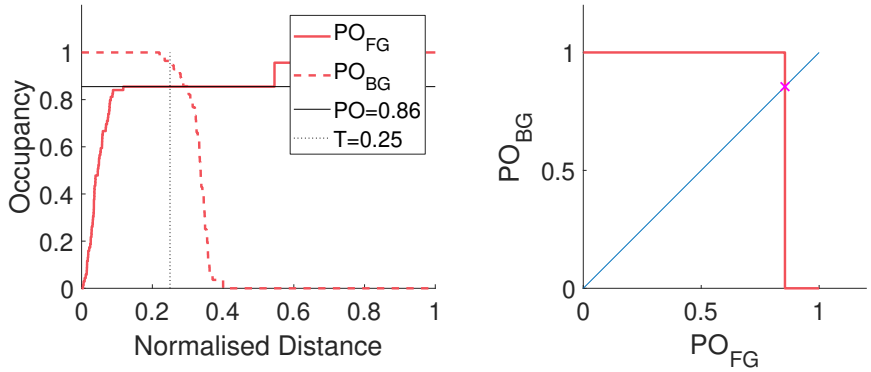
Figure B.3.10: Occupancy and PO plots used in detecting the amber lamps in images containing Bipolar Impulsive Noise from Figure B.3.1.



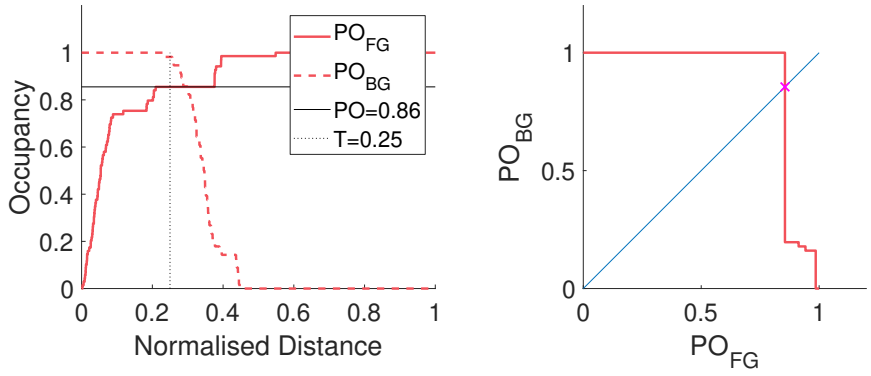
(a) $P(n) = 0.1,\ \rho = 0$



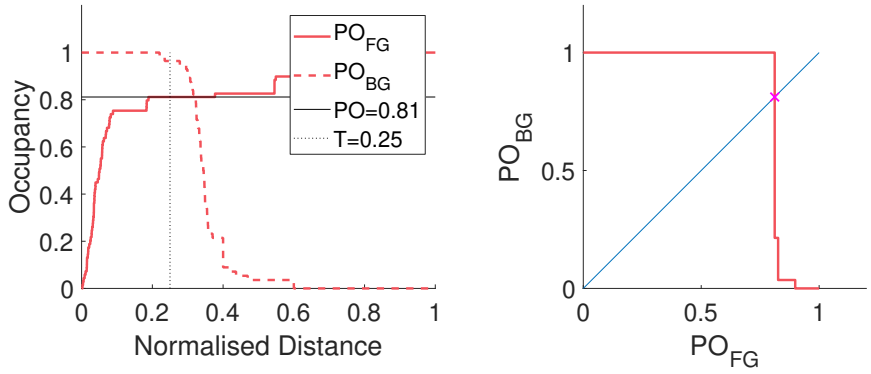
(b) $P(n) = 0.1,\ \rho = 0.5$



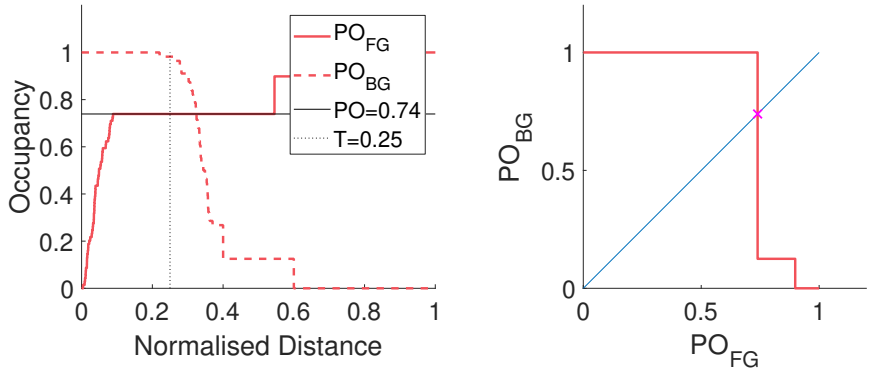
(c) $P(n) = 0.1,\ \rho = 1$



(d) $P(n) = 0.3,\ \rho = 0$



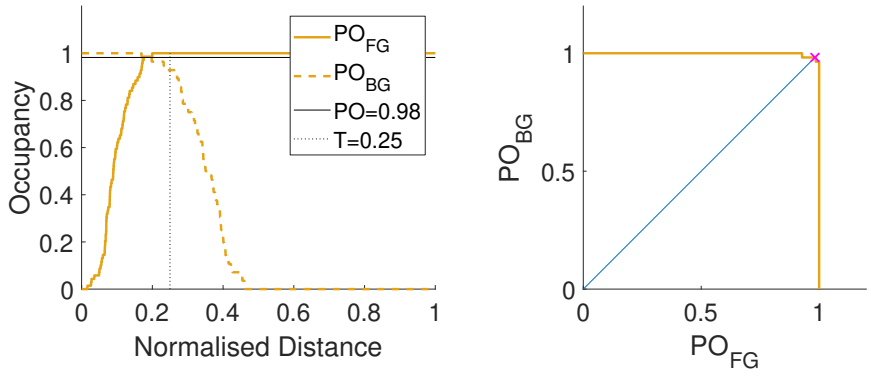
(e) $P(n) = 0.3,\ \rho = 0.5$



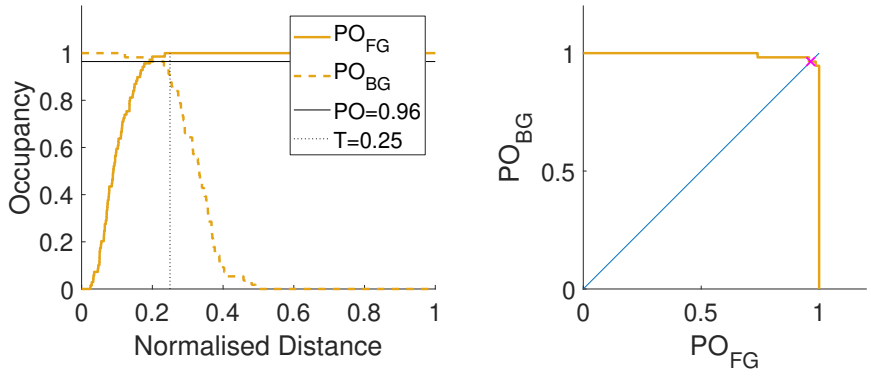
(f) $P(n) = 0.3,\ \rho = 1$

Figure B.3.11: Occupancy and PO plots used in detecting the red lamps in images containing Bipolar Impulsive Noise from Figure B.3.1.

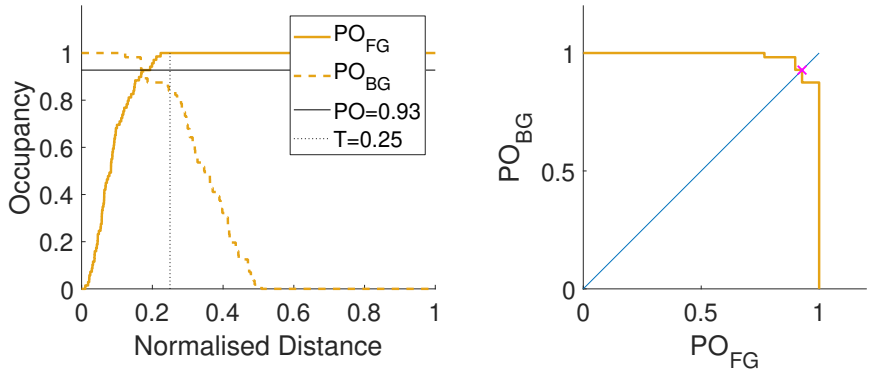
B.3.4.2 Gaussian Noise



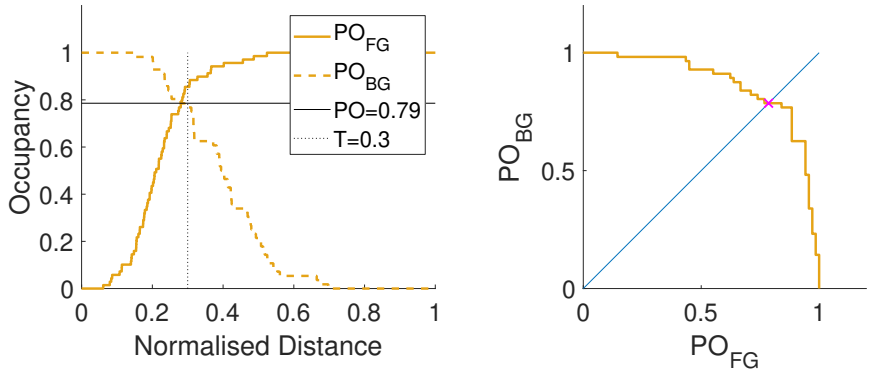
(a) $\sigma = 0.01,\ \rho = 0$



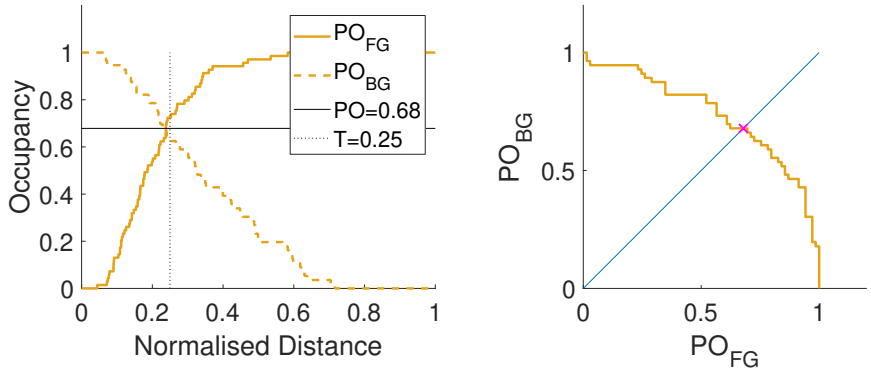
(b) $\sigma = 0.01,\ \rho = 0.5$



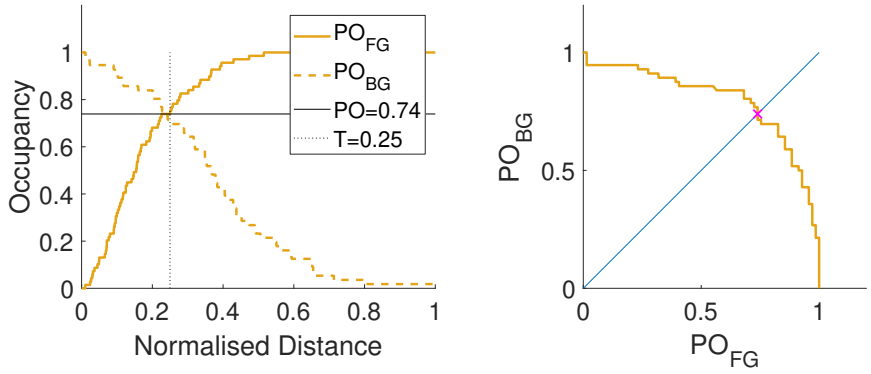
(c) $\sigma = 0.01,\ \rho = 1$



(d) $\sigma = 0.05,\ \rho = 0$

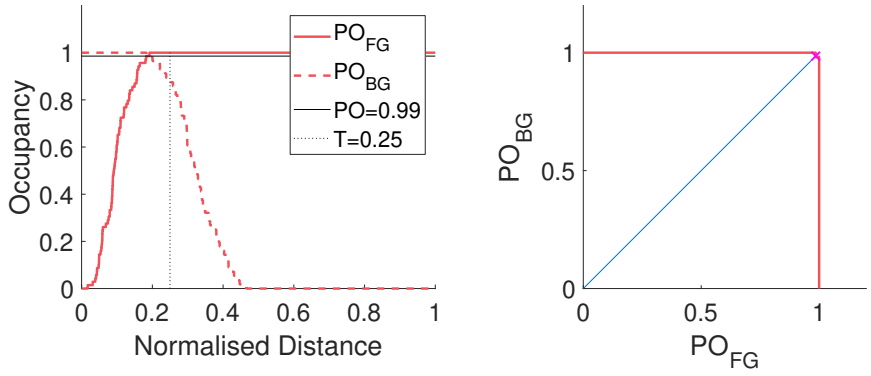


(e) $\sigma = 0.05,\ \rho = 0.5$

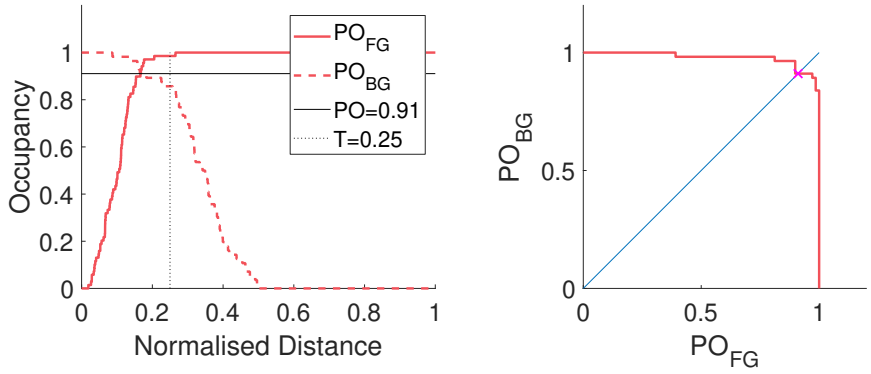


(f) $\sigma = 0.05,\ \rho = 1$

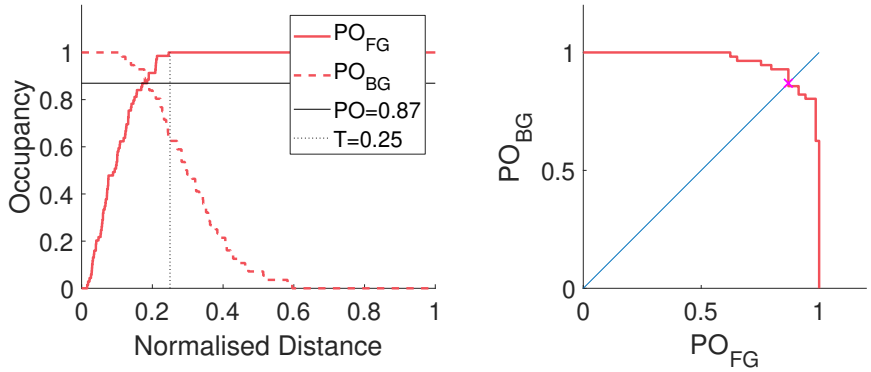
Figure B.3.12: Occupancy and PO plots used in detecting the amber lamps in images containing Gaussian Noise from Figure B.3.2.



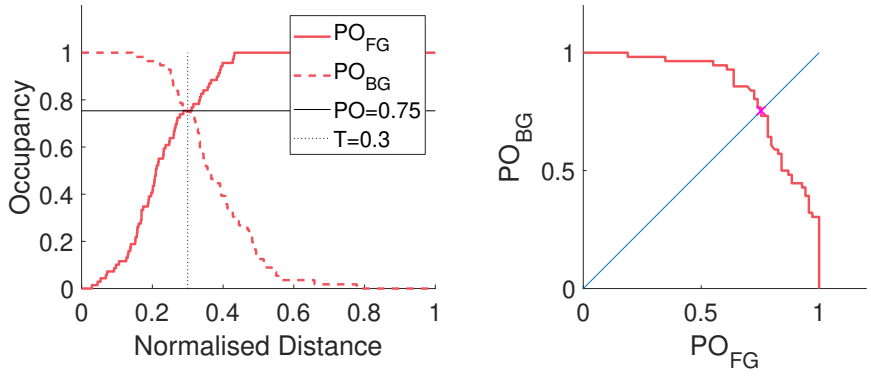
(a) $\sigma = 0.01,\ \rho = 0$



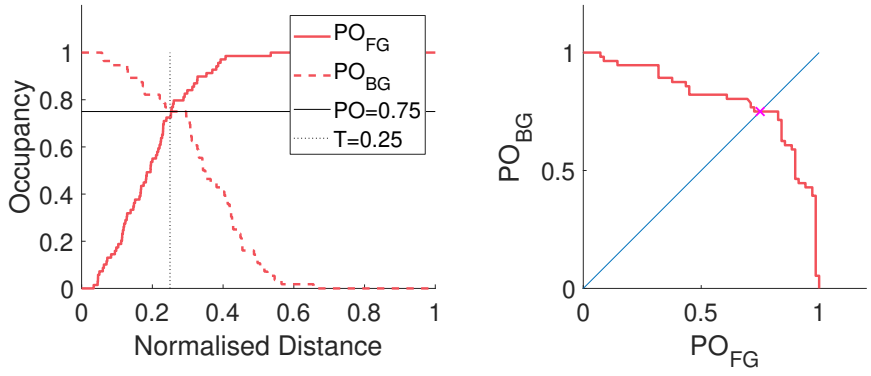
(b) $\sigma = 0.01,\ \rho = 0.5$



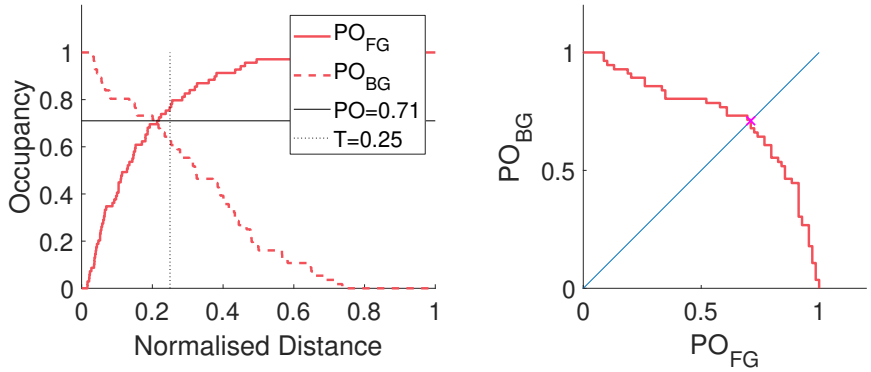
(c) $\sigma = 0.01,\ \rho = 1$



(d) $\sigma = 0.05,\ \rho = 0$



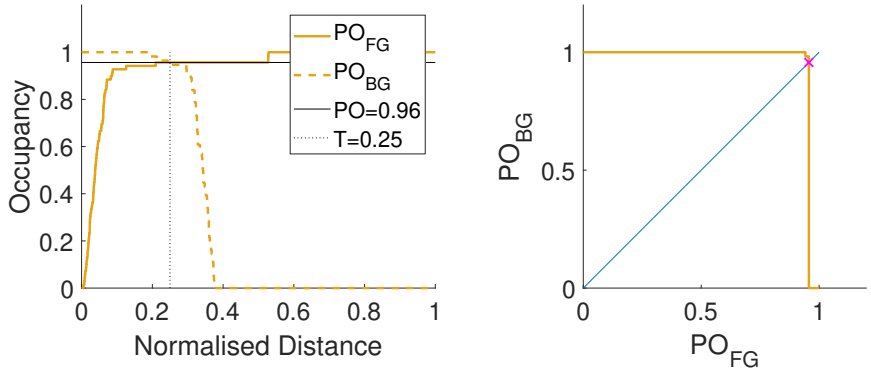
(e) $\sigma = 0.05,\ \rho = 0.5$



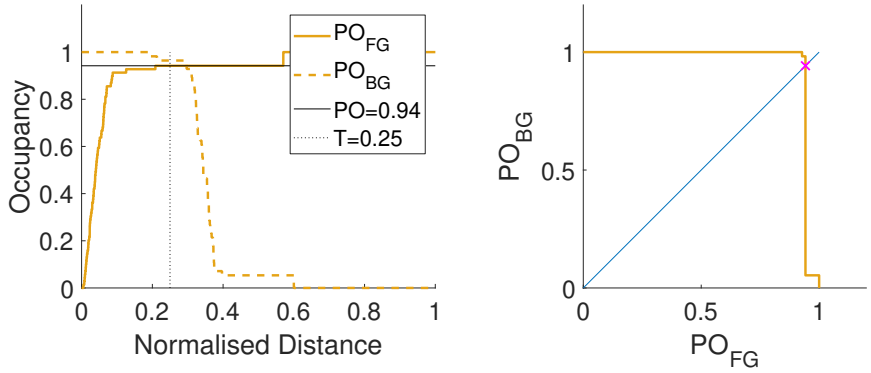
(f) $\sigma = 0.05,\ \rho = 1$

Figure B.3.13: Occupancy and PO plots used in detecting the red lamps in images containing Gaussian Noise from Figure B.3.2.

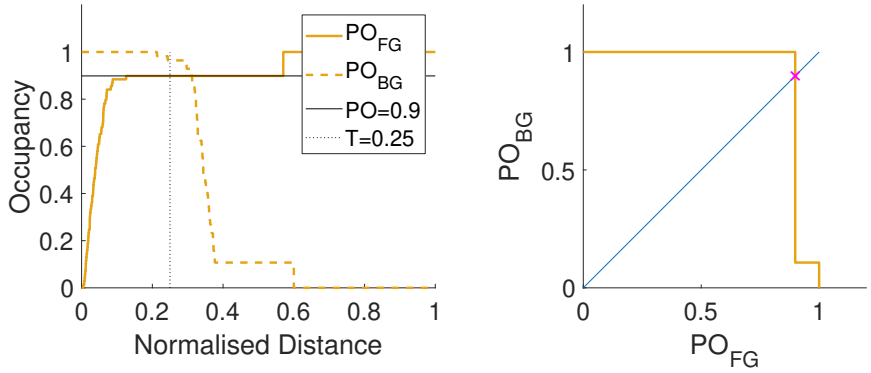
B.3.5 Impulsive Noise



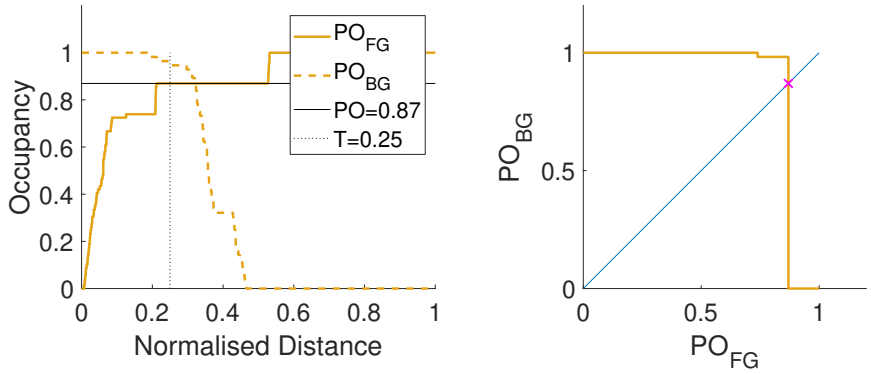
(a) $P(n) = 0.1,\ \rho = 0$



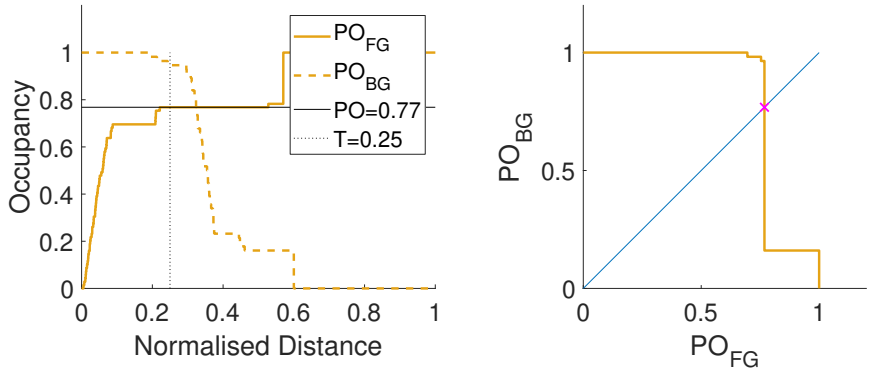
(b) $P(n) = 0.1,\ \rho = 0.5$



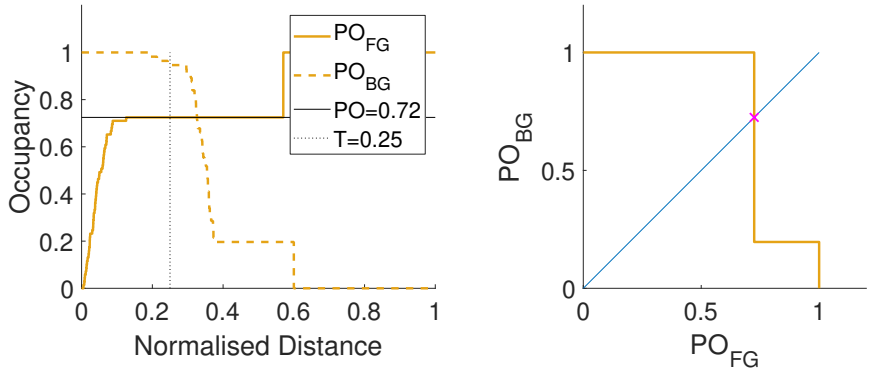
(c) $P(n) = 0.1,\ \rho = 1$



(d) $P(n) = 0.3,\ \rho = 0$

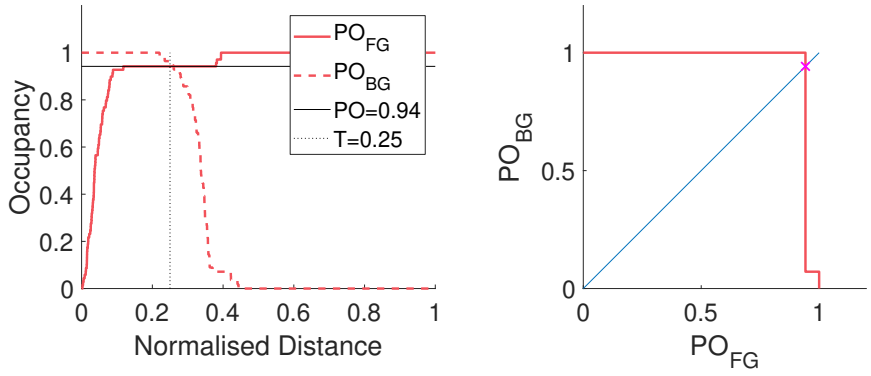


(e) $P(n) = 0.3,\ \rho = 0.5$

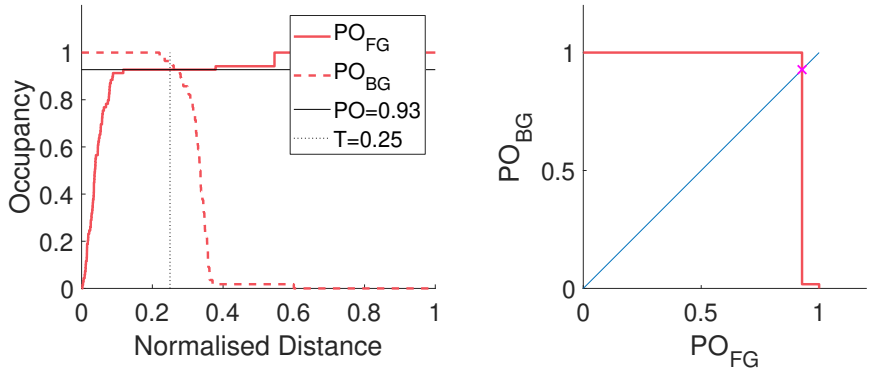


(f) $P(n) = 0.3,\ \rho = 1$

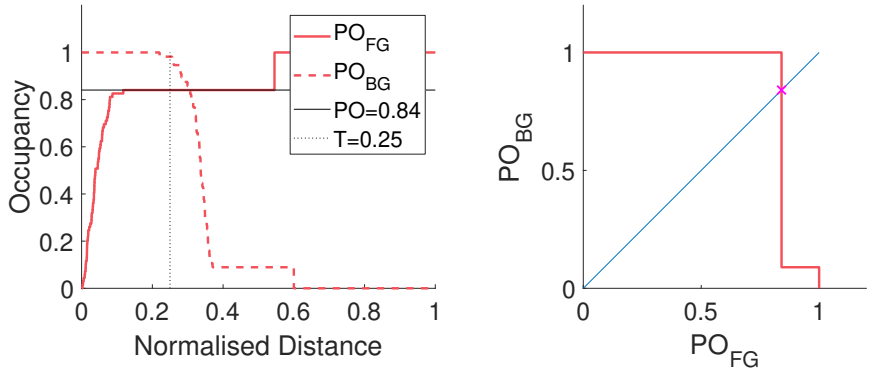
Figure B.3.14: Occupancy and PO plots used in detecting the amber lamps in images containing Impulsive Noise from Figure B.3.3.



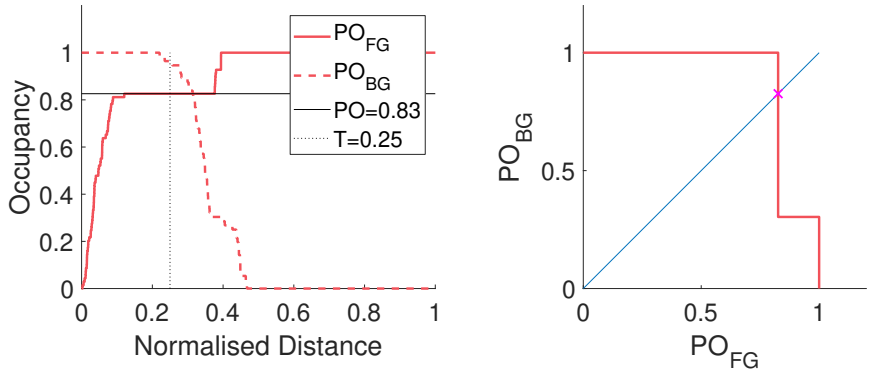
(a) $P(n) = 0.1,\ \rho = 0$



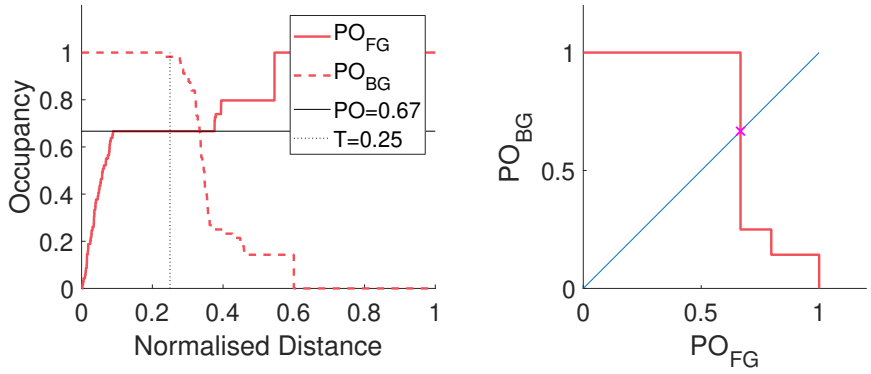
(b) $P(n) = 0.1,\ \rho = 0.5$



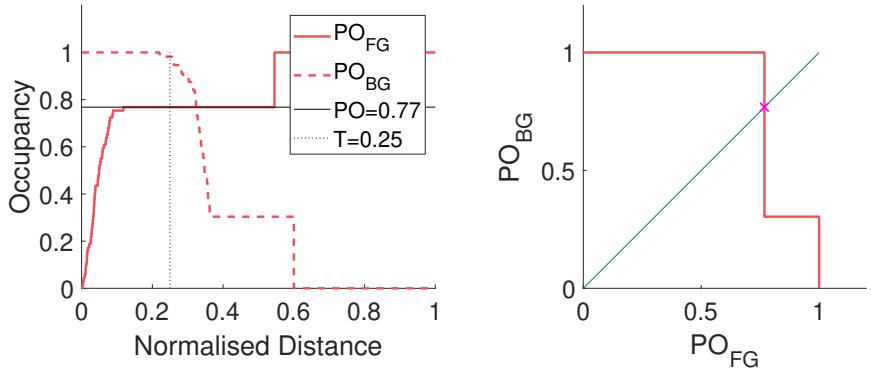
(c) $P(n) = 0.1,\ \rho = 1$



(d) $P(n) = 0.3,\ \rho = 0$



(e) $P(n) = 0.3,\ \rho = 0.5$



(f) $P(n) = 0.3,\ \rho = 1$

Figure B.3.15: Occupancy and PO plots used in detecting the red lamps in images containing Impulsive Noise from Figure B.3.3.

B.4 MDPOHMT Detections using the DOTA dataset

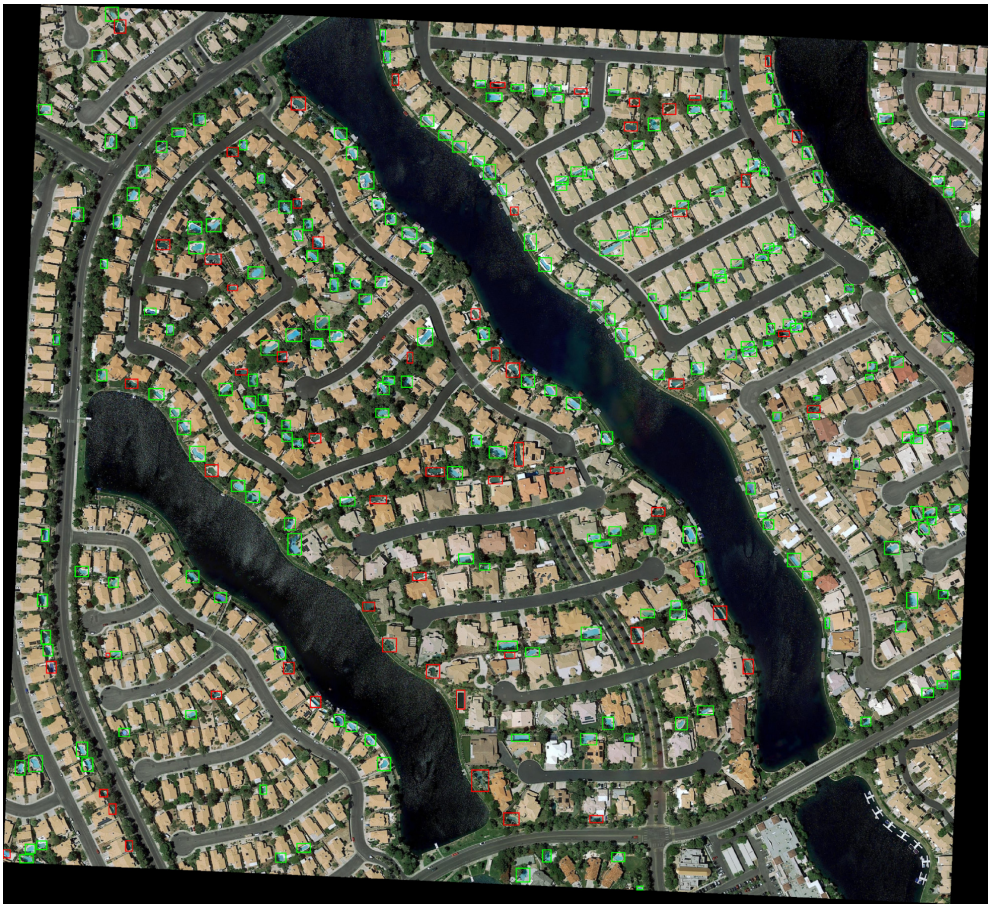


Figure B.4.1: MDPOHMT result on Figure A.1.1 with $p = 80\%$ and $T = 0.15$.

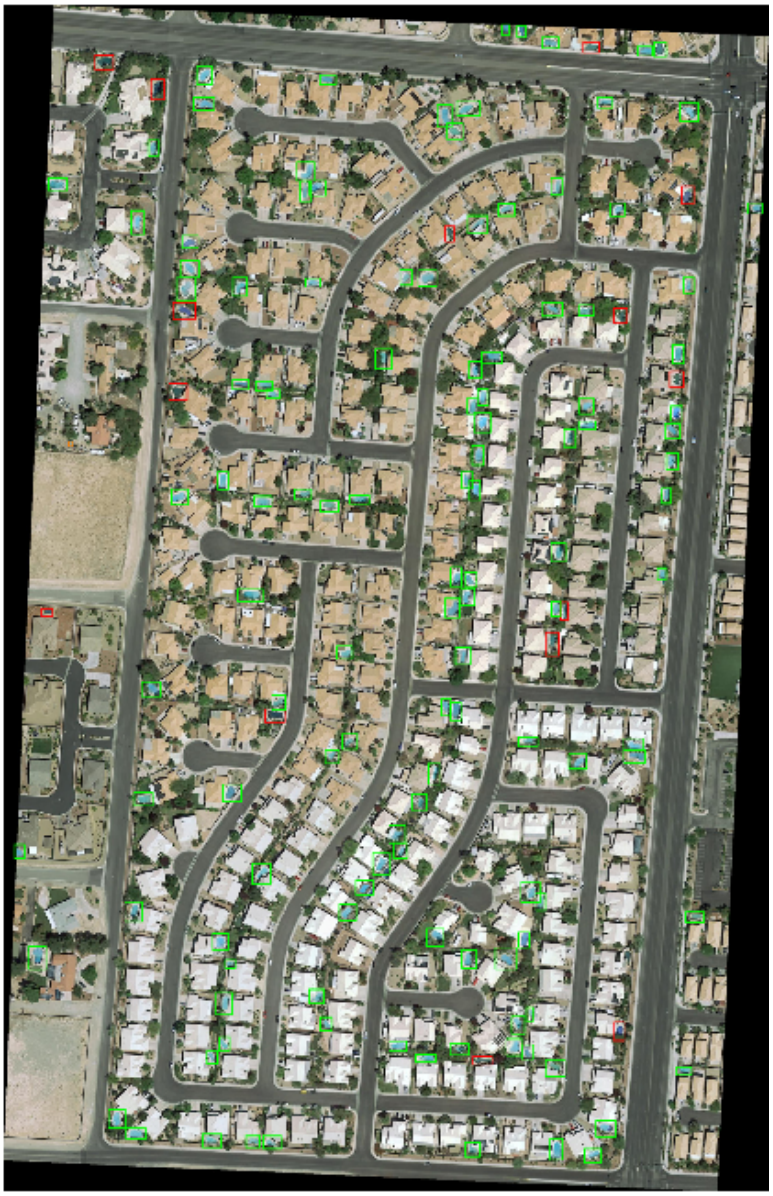


Figure B.4.2: MDPOHMT result on Figure A.1.2 with $p = 80\%$ and $T = 0.15$.

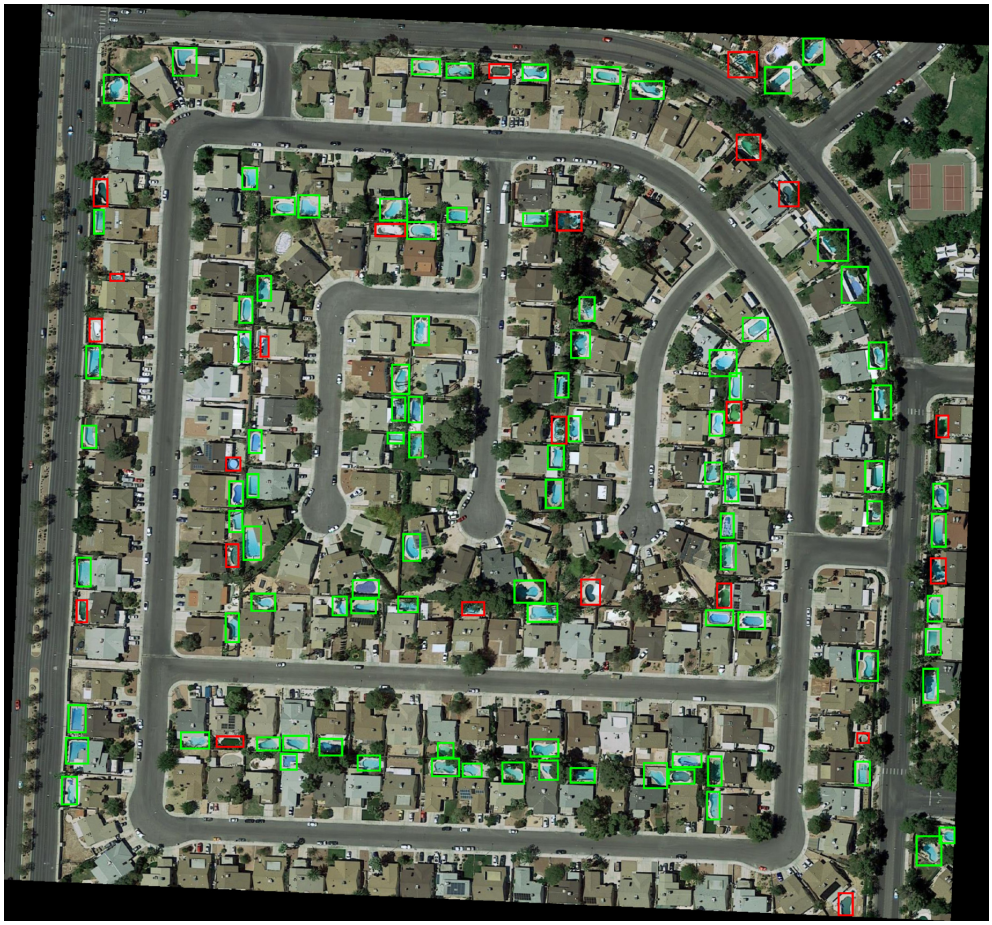


Figure B.4.3: MDPOHMT result on Figure A.1.3 with $p = 80\%$ and $T = 0.15$.

Figure B.4.4: MDPOHMT result on Figure A.1.4 with $p = 80\%$ and $T = 0.15$.

Figure B.4.5: MDPOHMT result on Figure A.1.5 with $p = 80\%$ and $T = 0.15$.

Figure B.4.6: MDPOHMT result on Figure A.1.6 with $p = 80\%$ and $T = 0.15$.

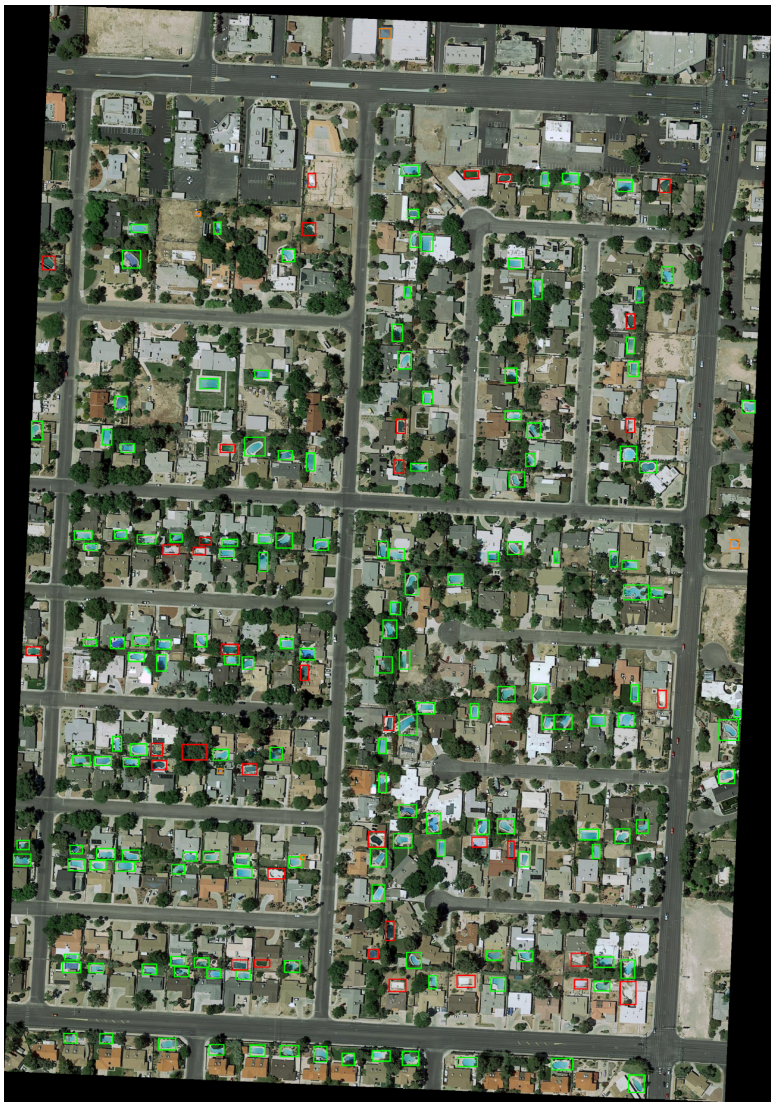


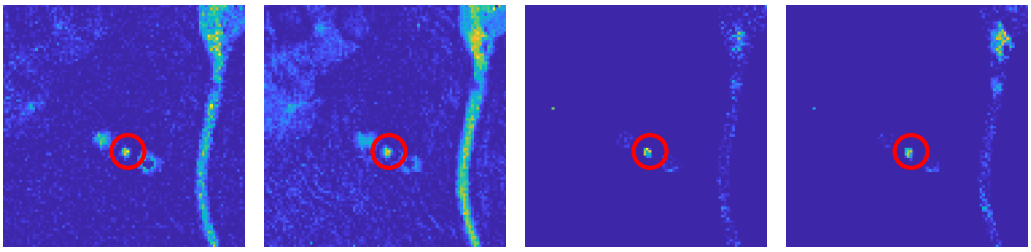
Figure B.4.7: MDPOHMT result on Figure A.1.7 with $p = 80\%$ and $T = 0.15$.

B.5 Additional Results using JSSDR-TD on the OP7 Dataset

B.5.1 OP7 1

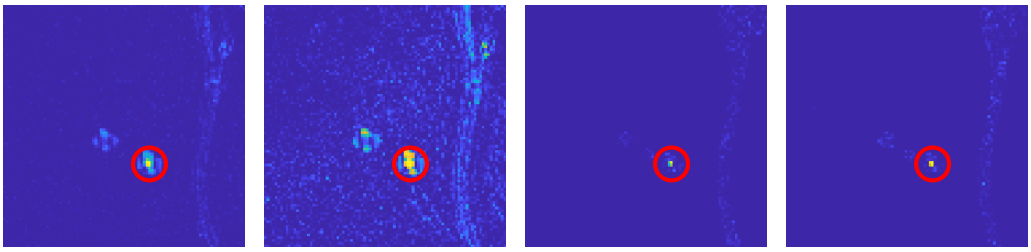
(a) (b)

Figure B.5.1: Target region of the OP7 1 image **a)** before, and **b)** after NDVI_{re}-based spatial masking.



(a) (b) (c) (d)

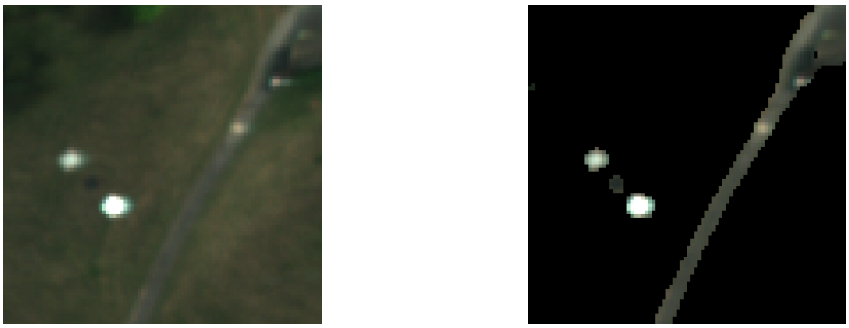
Figure B.5.2: JSSDR-TD results detecting target 2 (black tile) in the OP7 1 Image with **a)** Raw ($L = 100$) data, **b)** PCA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** PCA and NDVI_{re}.



(a) (b) (c) (d)

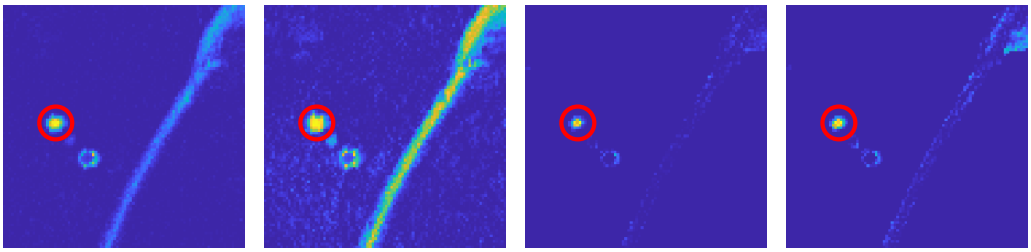
Figure B.5.3: JSSDR-TD results detecting target 3 (white tile) in the OP7 1 Image with **a)** Raw ($L = 100$) data, **b)** PCA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** PCA and NDVI_{re}.

B.5.2 OP7 2



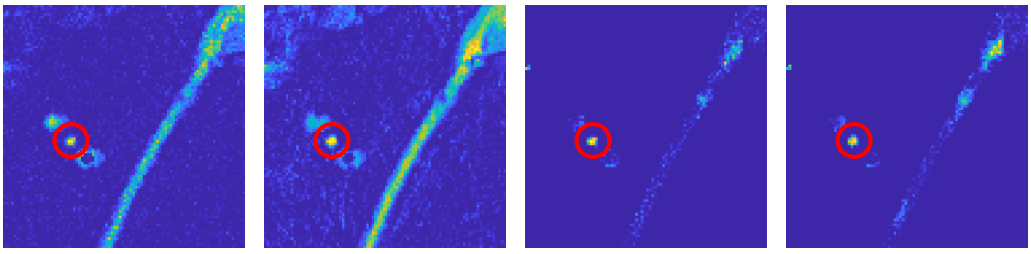
(a) (b)

Figure B.5.4: Target region of the OP7 2 image **a)** before, and **b)** after NDVI_{re}-based spatial masking.



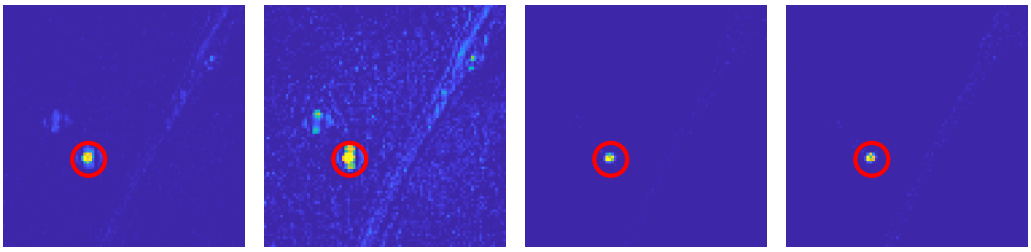
(a) (b) (c) (d)

Figure B.5.5: JSSDR-TD results detecting target 1 (grey tile) in the OP7 2 Image with **a)** Raw ($L = 100$) data, **b)** PCA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** PCA and NDVI_{re}.



(a) (b) (c) (d)

Figure B.5.6: JSSDR-TD results detecting target 2 (black tile) in the OP7 2 Image with **a)** Raw ($L = 100$) data, **b)** PCA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** PCA and NDVI_{re}.



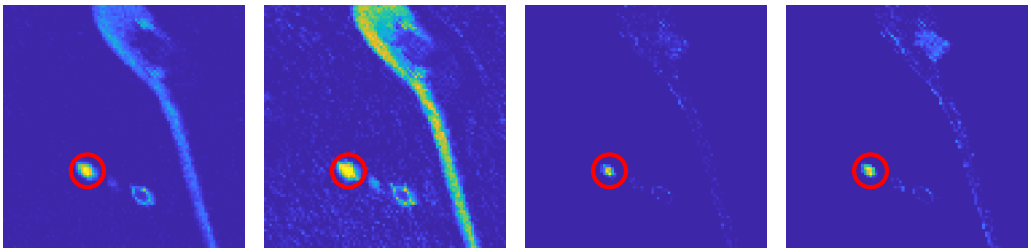
(a) (b) (c) (d)

Figure B.5.7: JSSDR-TD results detecting target 3 (white tile) in the OP7 2 Image with **a)** Raw ($L = 100$) data, **b)** PCA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** PCA and NDVI_{re}.

B.5.3 OP7 3

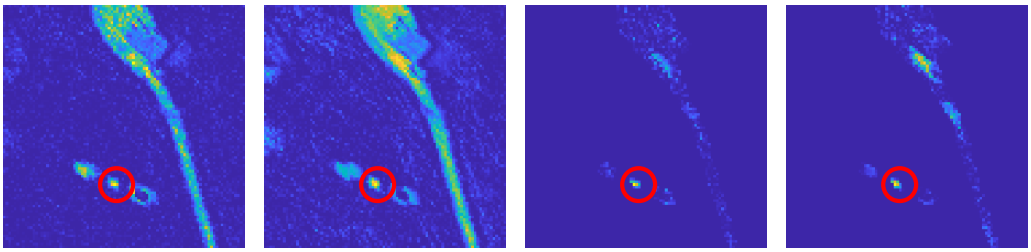
(a) (b)

Figure B.5.8: Target region of the OP7 3 image **a)** before, and **b)** after NDVI_{re}-based spatial masking.



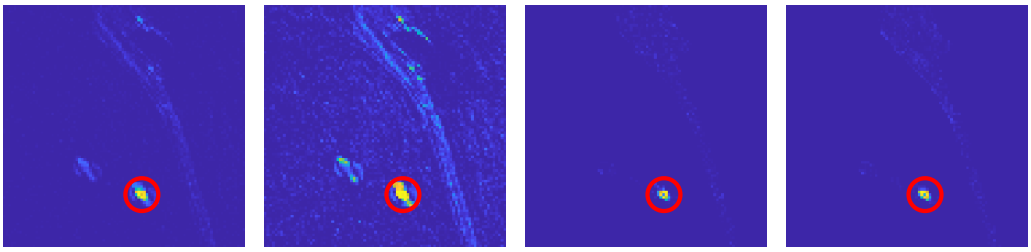
(a) (b) (c) (d)

Figure B.5.9: JSSDR-TD results detecting target 1 (grey tile) in the OP7 3 Image with **a)** Raw ($L = 100$) data, **b)** PCA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** PCA and NDVI_{re}.



(a) (b) (c) (d)

Figure B.5.10: JSSDR-TD results detecting target 2 (black tile) in the OP7 3 Image with **a)** Raw ($L = 100$) data, **b)** PCA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** PCA and NDVI_{re}.



(a) (b) (c) (d)

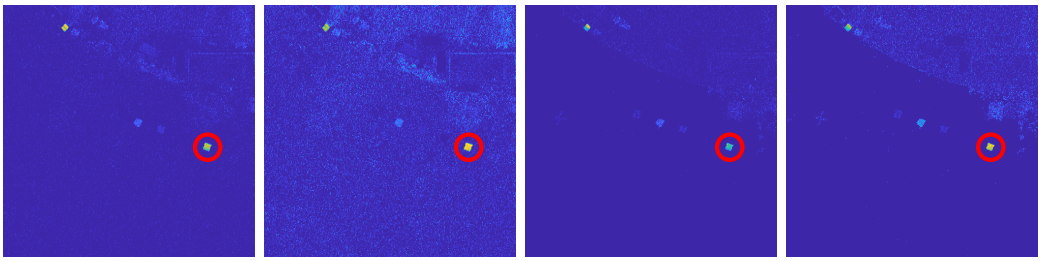
Figure B.5.11: JSSDR-TD results detecting target 3 (white tile) in the OP7 3 Image with **a)** Raw ($L = 100$) data, **b)** PCA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** PCA and NDVI_{re}.

B.6 Additional Results using JSSDR-TD on the UDRC Dataset

B.6.1 IM140804

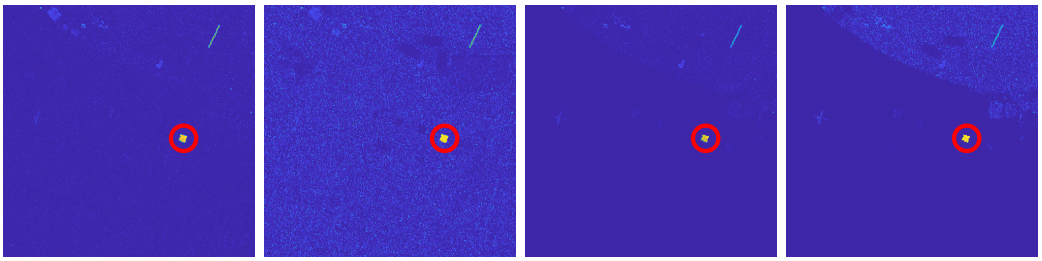
(a) (b)

Figure B.6.1: Target region of the IM 140804 image **a)** before and **b)** after NDVI_{re}-based spatial masking.



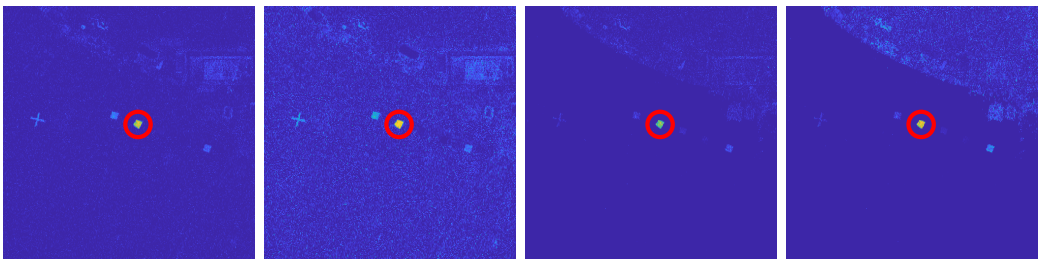
(a) (b) (c) (d)

Figure B.6.2: JSSDR-TD results detecting target 1 (brown carpet) in the IM140804 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



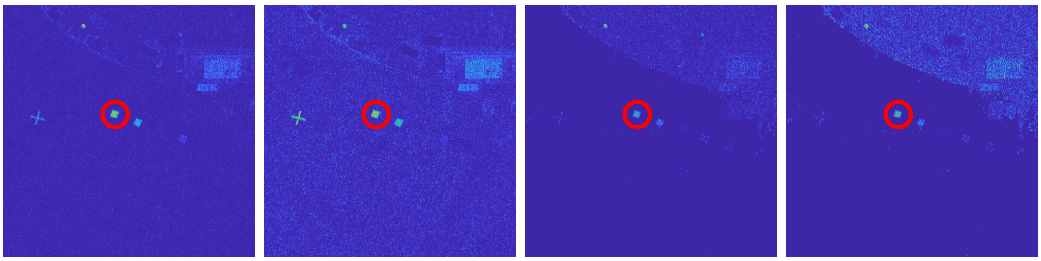
(a) (b) (c) (d)

Figure B.6.3: JSSDR-TD results detecting target 2 (green carpet) in the IM140804 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

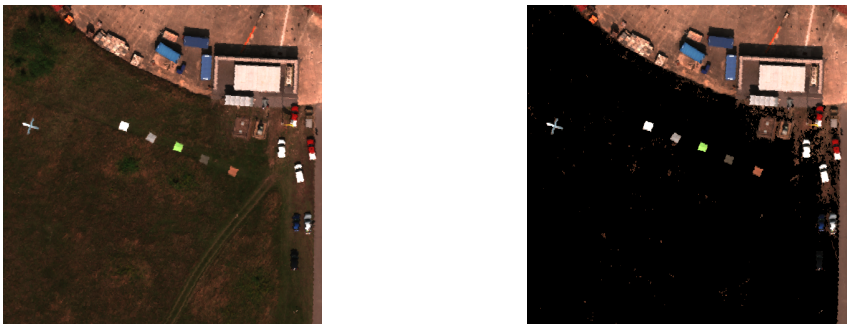
Figure B.6.4: JSSDR-TD results detecting target 5 (grey ceramic) in the IM140804 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

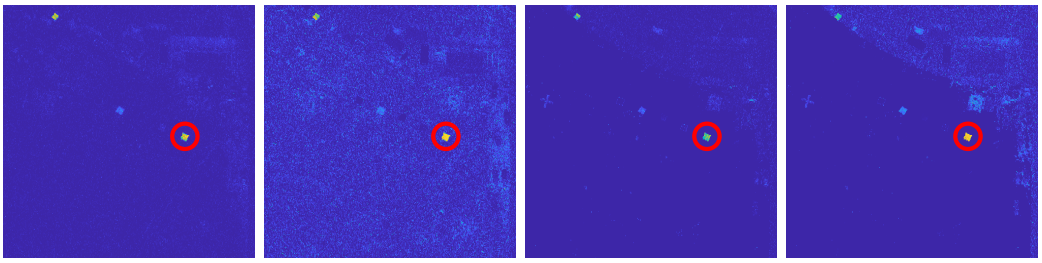
Figure B.6.5: JSSDR-TD results detecting target 7 (white perspex) in the IM140804 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.

B.6.2 IM140806



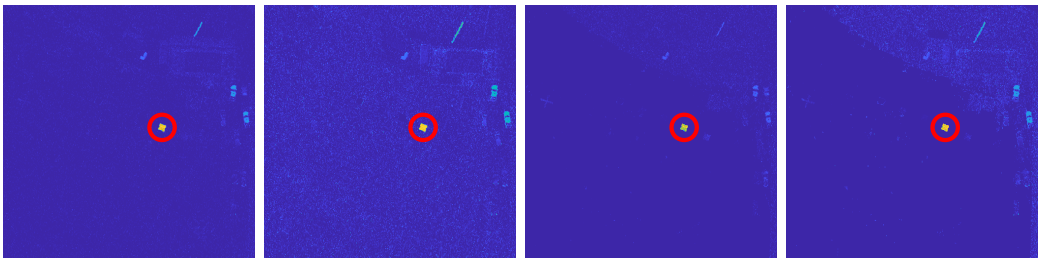
(a) (b)

Figure B.6.6: Target region of the IM 140806 image **a)** before and **b)** after NDVI_{re}-based spatial masking.



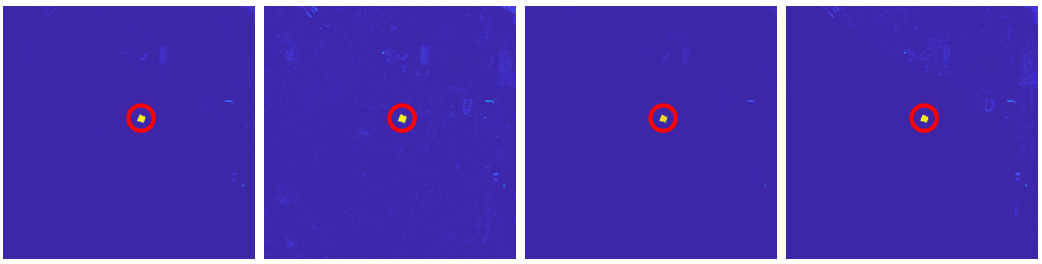
(a) (b) (c) (d)

Figure B.6.7: JSSDR-TD results detecting target 1 (brown carpet) in the IM140806 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



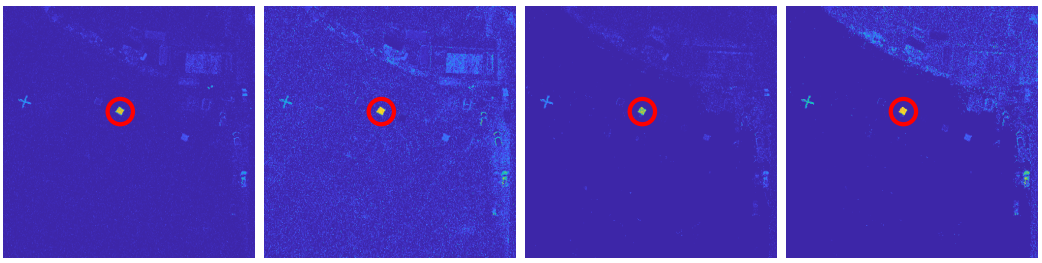
(a) (b) (c) (d)

Figure B.6.8: JSSDR-TD results detecting target 2 (green carpet) in the IM140806 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



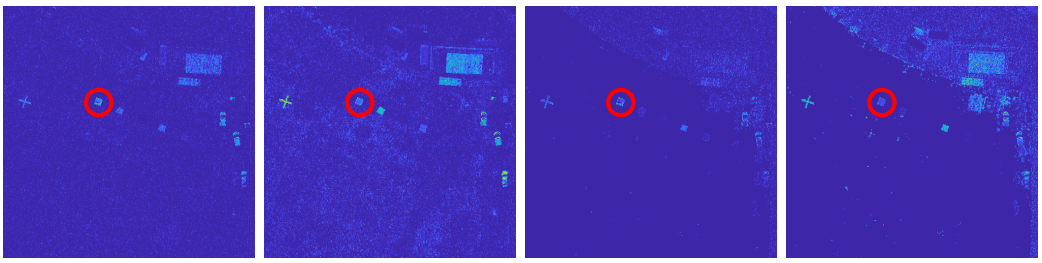
(a) (b) (c) (d)

Figure B.6.9: JSSDR-TD results detecting target 3 (green ceramic) in the IM140806 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

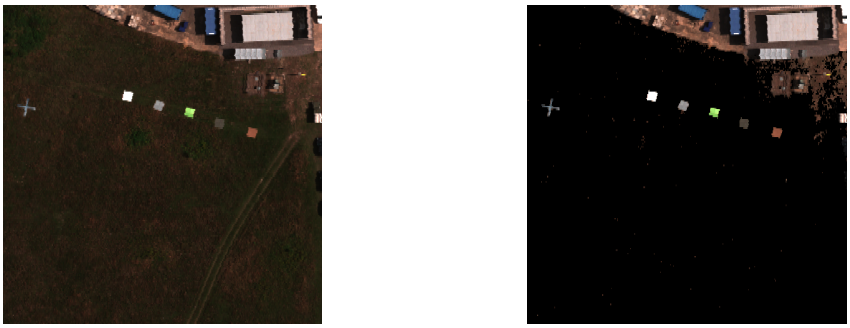
Figure B.6.10: JSSDR-TD results detecting target 5 (grey ceramic) in the IM140806 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

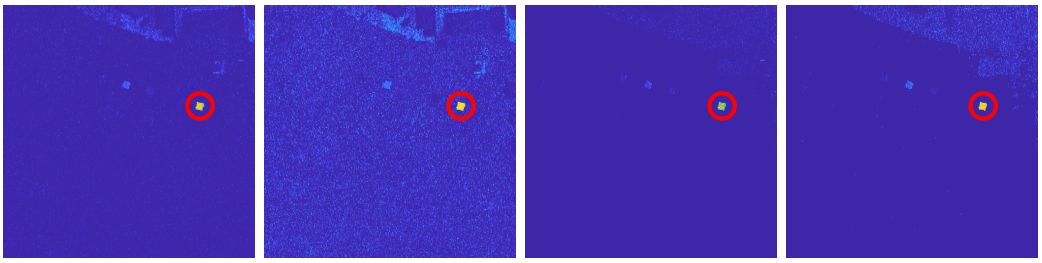
Figure B.6.11: JSSDR-TD results detecting target 7 (white perspex) in the IM140806 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.

B.6.3 IM140807



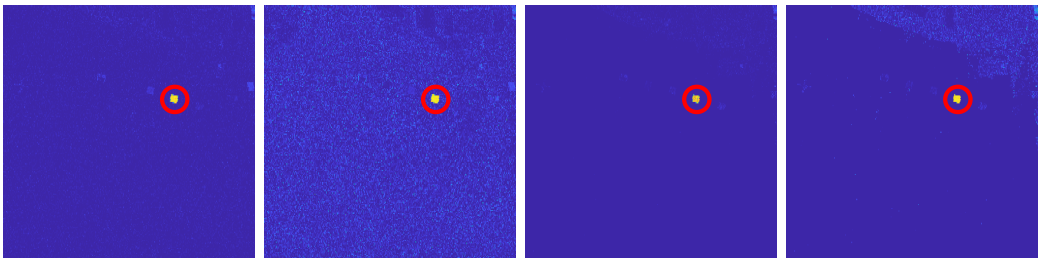
(a) (b)

Figure B.6.12: Target region of the IM 140807 image **a)** before and **b)** after NDVI_{re}-based spatial masking.



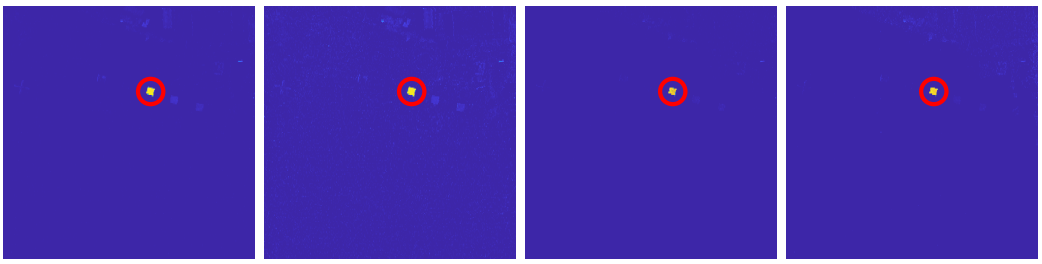
(a) (b) (c) (d)

Figure B.6.13: JSSDR-TD results detecting target 1 (brown carpet) in the IM140807 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



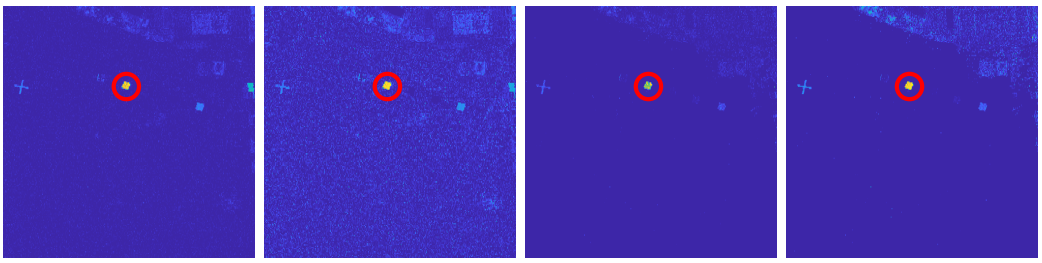
(a) (b) (c) (d)

Figure B.6.14: JSSDR-TD results detecting target 2 (green carpet) in the IM140807 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



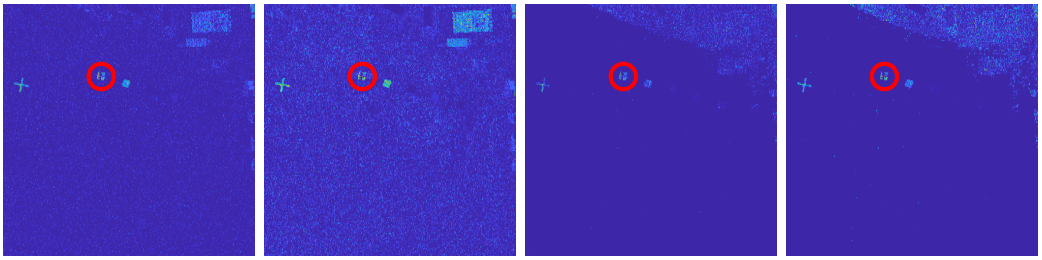
(a) (b) (c) (d)

Figure B.6.15: JSSDR-TD results detecting target 3 (green ceramic) in the IM140807 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

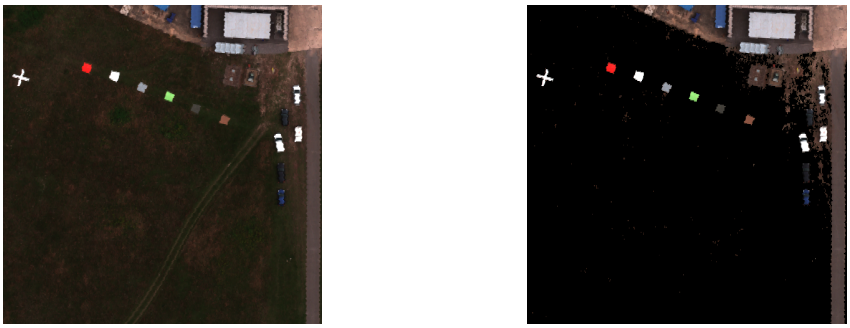
Figure B.6.16: JSSDR-TD results detecting target 5 (grey ceramic) in the IM140807 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

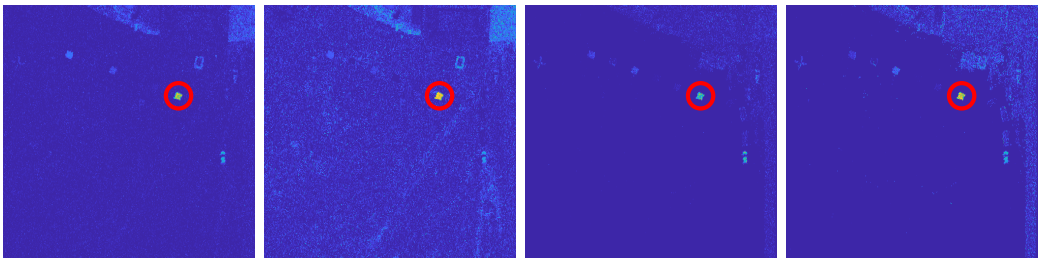
Figure B.6.17: JSSDR-TD results detecting target 7 (white perspex) in the IM140807 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.

B.6.4 IM140808



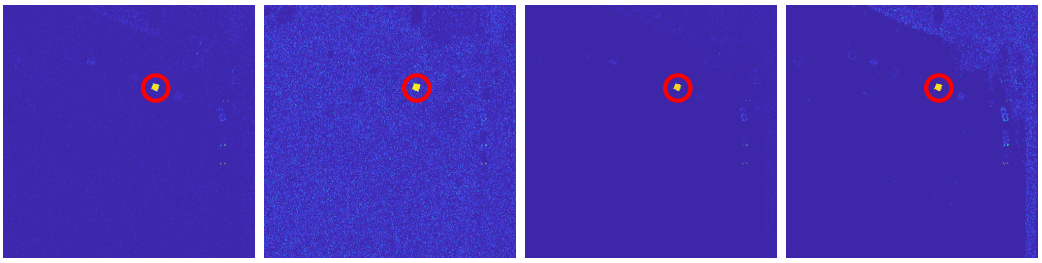
(a) (b)

Figure B.6.18: Target region of the IM 140808 image **a)** before and **b)** after NDVI_{re}-based spatial masking.



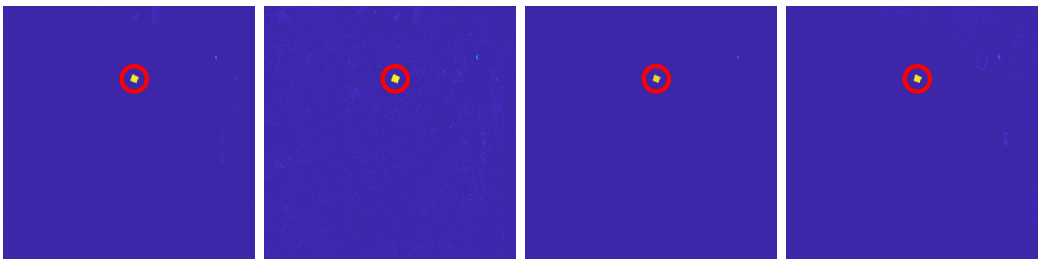
(a) (b) (c)

Figure B.6.19: JSSDR-TD results detecting target 1 (brown carpet) in the IM140808 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



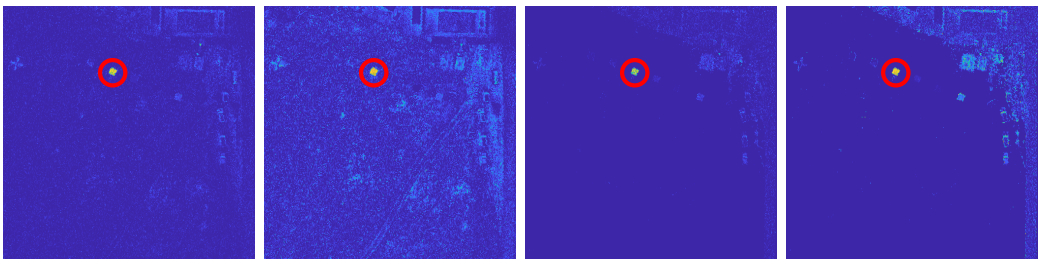
(a) (b) (c) (d)

Figure B.6.20: JSSDR-TD results detecting target 2 (green carpet) in the IM140808 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



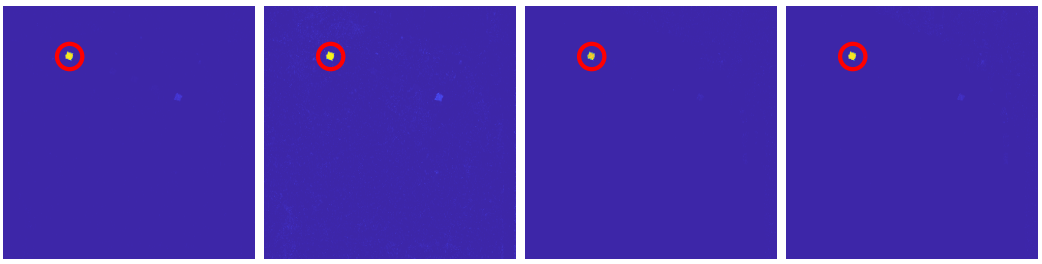
(a) (b) (c) (d)

Figure B.6.21: JSSDR-TD results detecting target 3 (green ceramic) in the IM140808 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



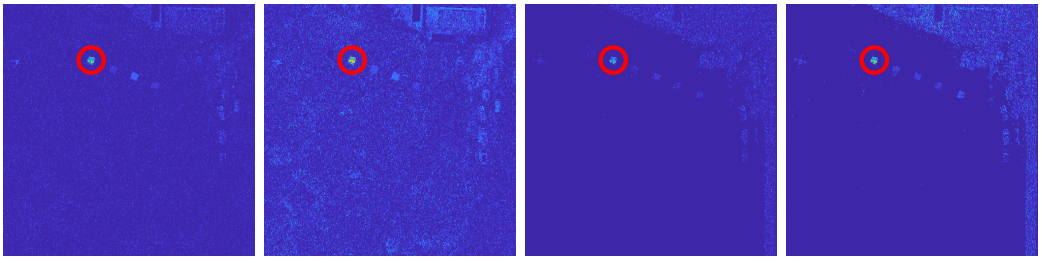
(a) (b) (c) (d)

Figure B.6.22: JSSDR-TD results detecting target 5 (grey ceramic) in the IM140808 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

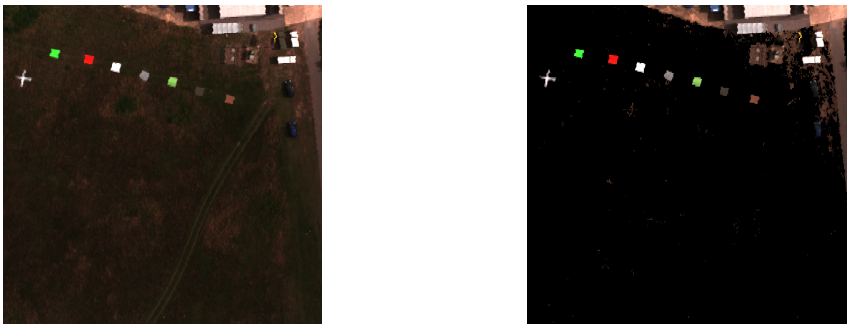
Figure B.6.23: JSSDR-TD results detecting target 6 (orange perspex) in the IM140808 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

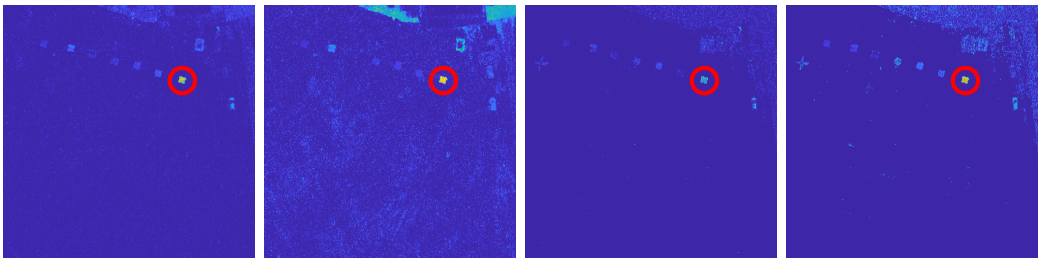
Figure B.6.24: JSSDR-TD results detecting target 7 (white perspex) in the IM140808 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.

B.6.5 IM140812



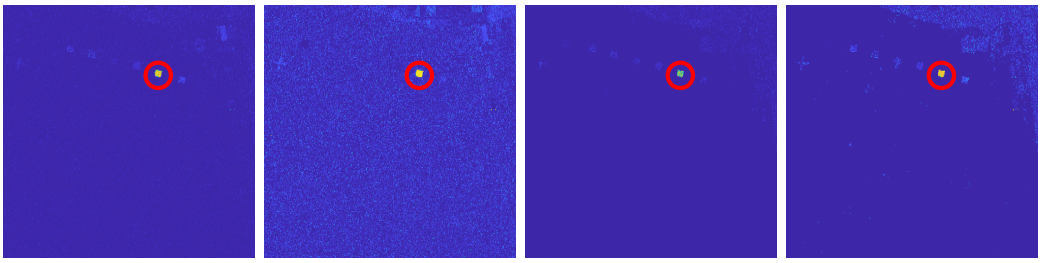
(a) (b)

Figure B.6.25: Target region of the IM 140812 image **a)** before and **b)** after NDVI_{re}-based spatial masking.



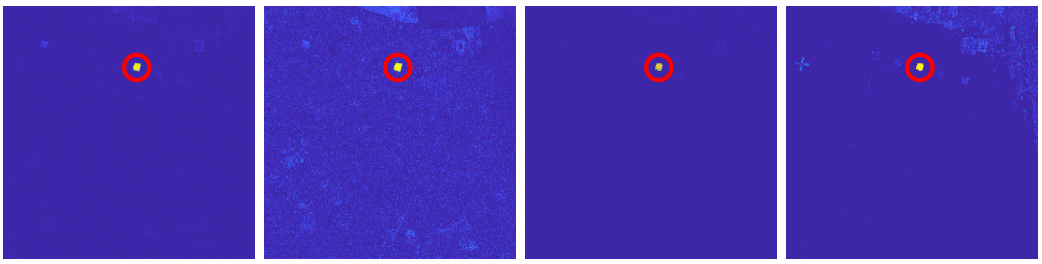
(a) (b) (c) (d)

Figure B.6.26: JSSDR-TD results detecting target 1 (brown carpet) in the IM140812 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



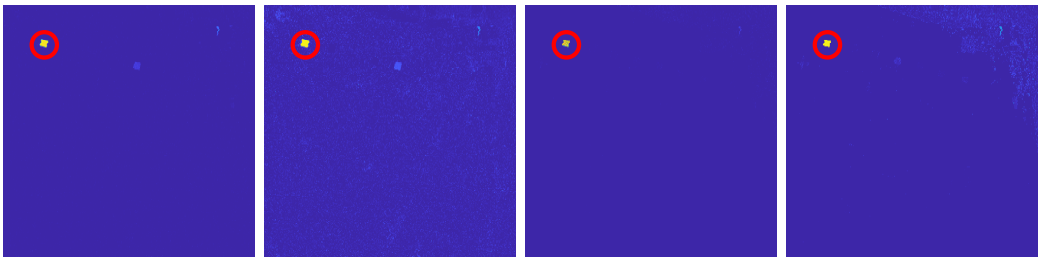
(a) (b) (c) (d)

Figure B.6.27: JSSDR-TD results detecting target 2 (green carpet) in the IM140812 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



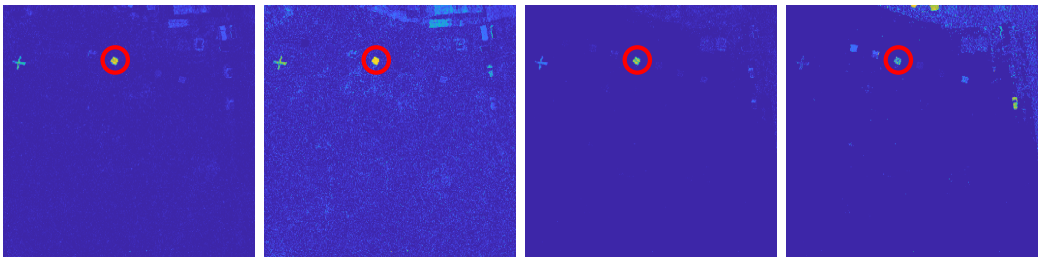
(a) (b) (c) (d)

Figure B.6.28: JSSDR-TD results detecting target 3 (green ceramic) in the IM140812 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



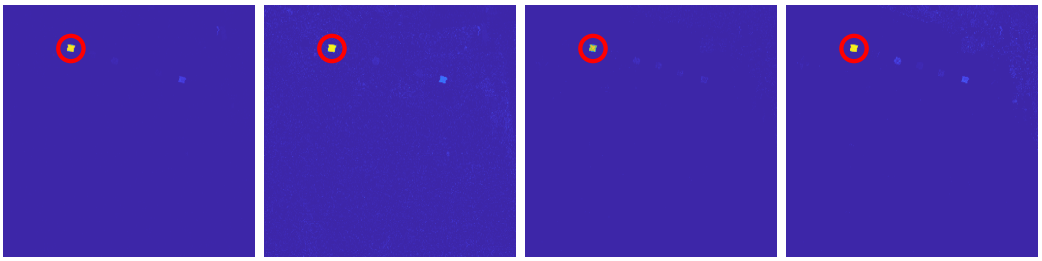
(a) (b) (c) (d)

Figure B.6.29: JSSDR-TD results detecting target 4 (green perspex) in the IM140812 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



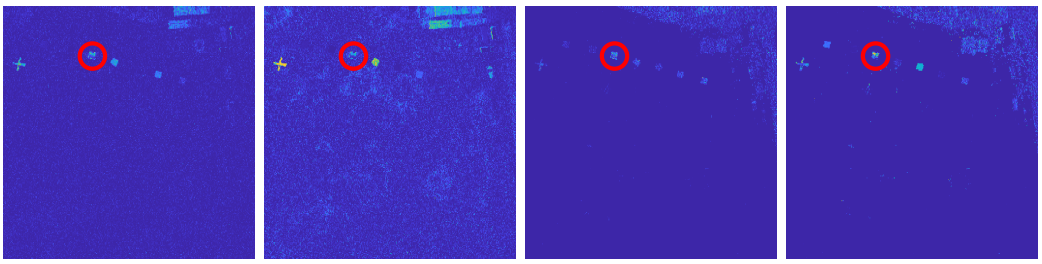
(a) (b) (c) (d)

Figure B.6.30: JSSDR-TD results detecting target 5 (grey ceramic) in the IM140812 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

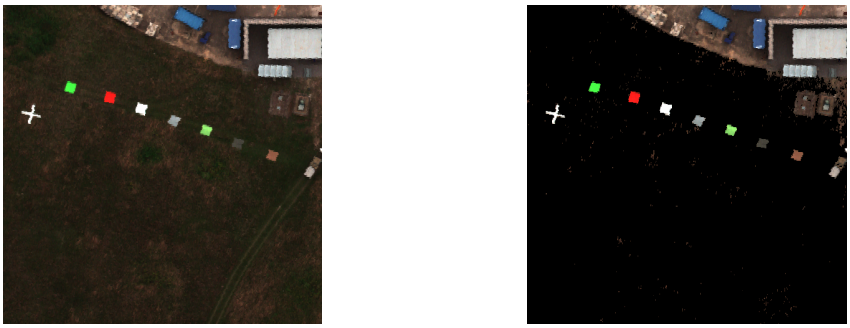
Figure B.6.31: JSSDR-TD results detecting target 6 (orange perspex) in the IM140812 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

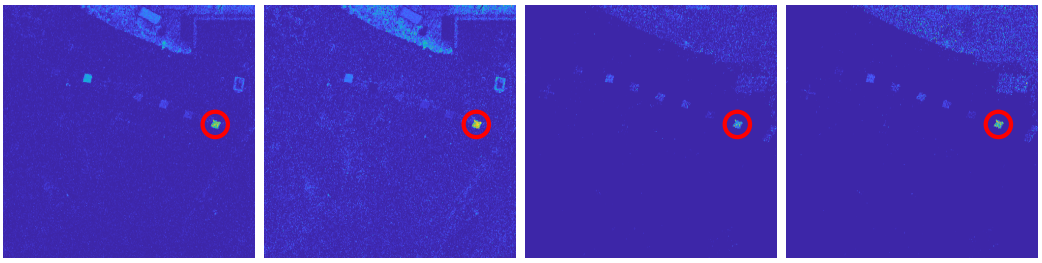
Figure B.6.32: JSSDR-TD results detecting target 7 (white perspex) in the IM140812 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.

B.6.6 IM140813



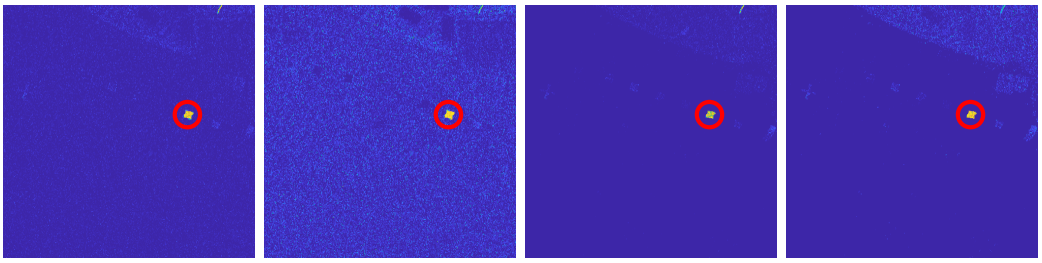
(a) (b)

Figure B.6.33: Target region of the IM 140813 image **a)** before and **b)** after NDVI_{re}-based spatial masking.



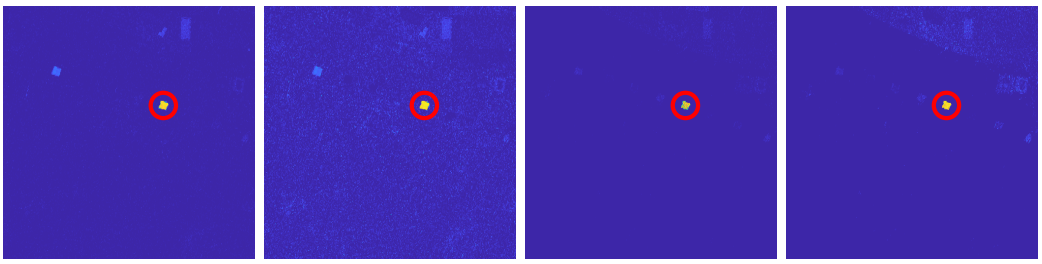
(a) (b) (c) (d)

Figure B.6.34: JSSDR-TD results detecting target 1 (brown carpet) in the IM140813 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



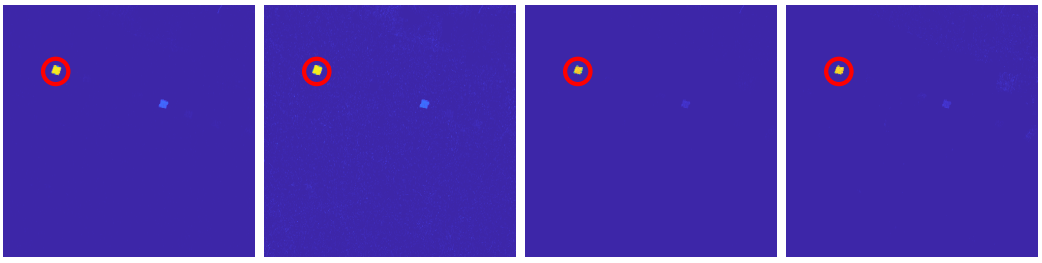
(a) (b) (c) (d)

Figure B.6.35: JSSDR-TD results detecting target 2 (green carpet) in the IM140813 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



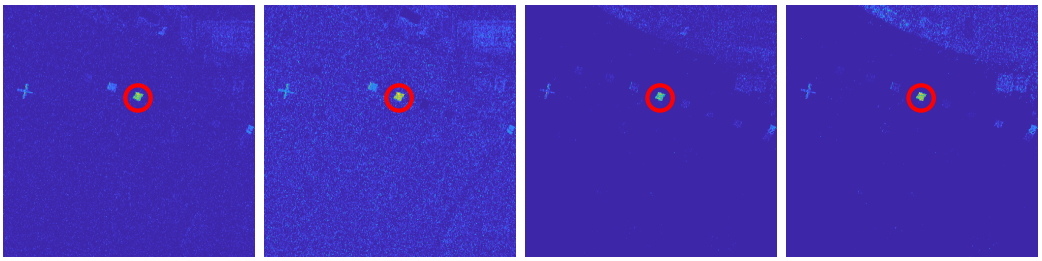
(a) (b) (c) (d)

Figure B.6.36: JSSDR-TD results detecting target 3 (green ceramic) in the IM140813 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



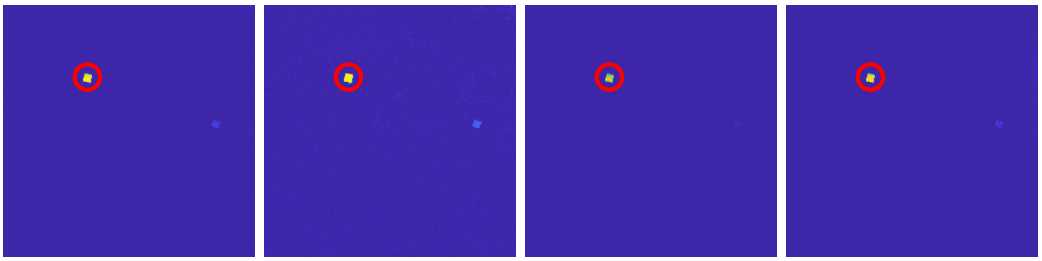
(a) (b) (c) (d)

Figure B.6.37: JSSDR-TD results detecting target 4 (green perspex) in the IM140813 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



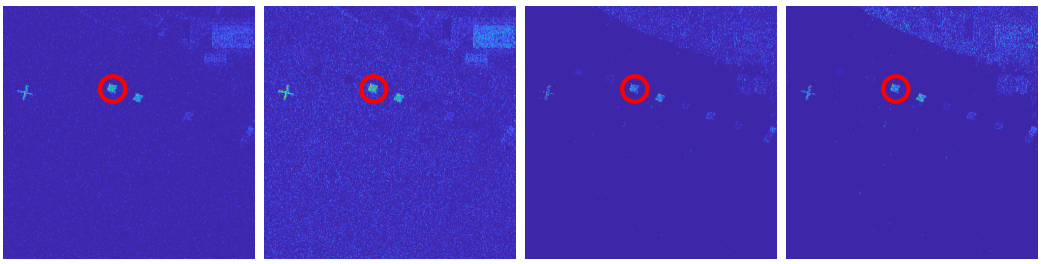
(a) (b) (c) (d)

Figure B.6.38: JSSDR-TD results detecting target 5 (grey ceramic) in the IM140813 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

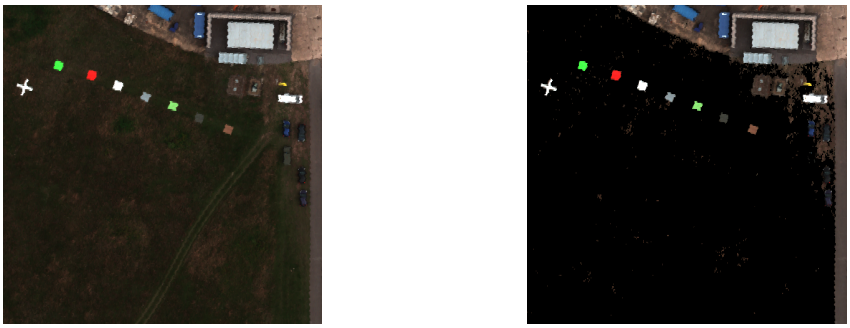
Figure B.6.39: JSSDR-TD results detecting target 6 (orange perspex) in the IM140813 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

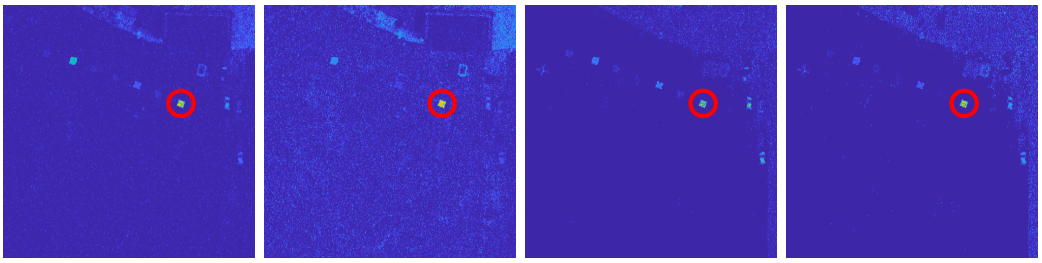
Figure B.6.40: JSSDR-TD results detecting target 7 (white perspex) in the IM140813 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.

B.6.7 IM140815



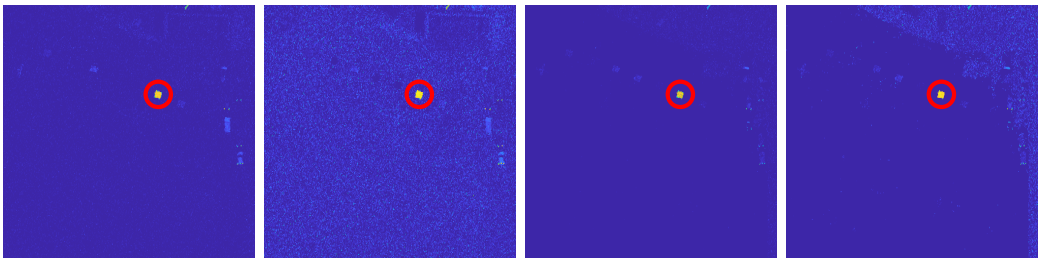
(a) (b)

Figure B.6.41: Target region of the IM 140815 image **a)** before and **b)** after NDVI_{re}-based spatial masking.



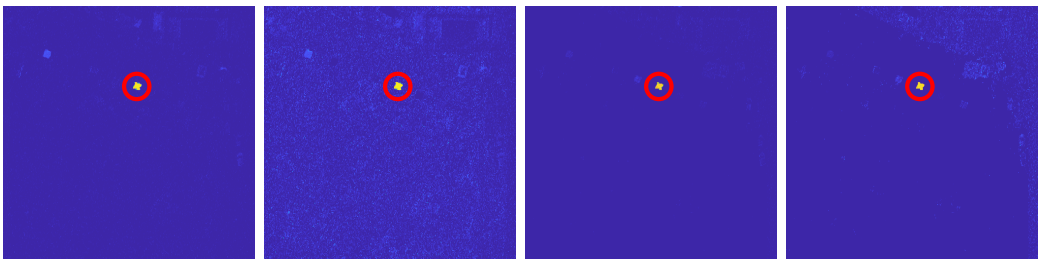
(a) (b) (c) (d)

Figure B.6.42: JSSDR-TD results detecting target 1 (brown carpet) in the IM140815 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



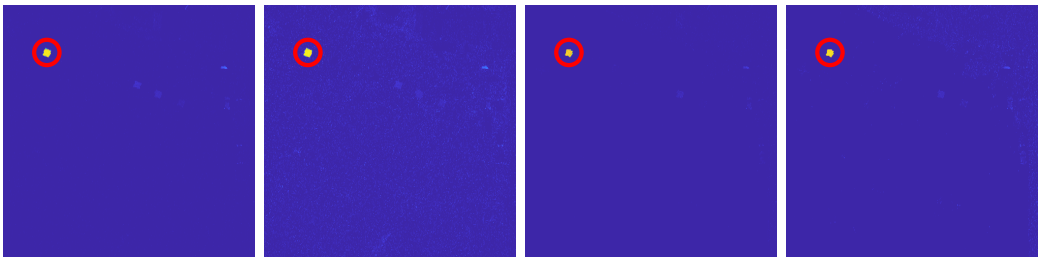
(a) (b) (c) (d)

Figure B.6.43: JSSDR-TD results detecting target 2 (green carpet) in the IM140815 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



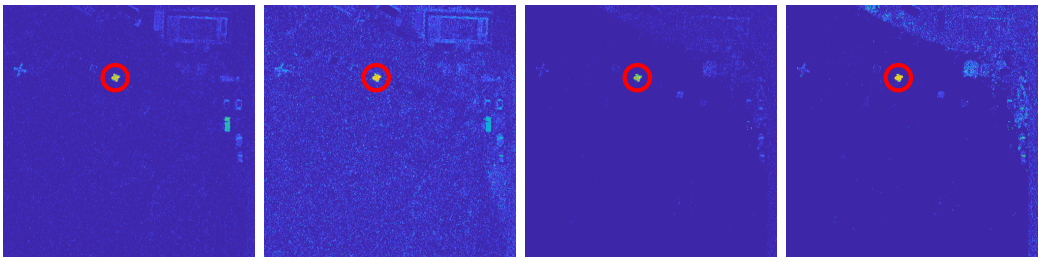
(a) (b) (c) (d)

Figure B.6.44: JSSDR-TD results detecting target 3 (green ceramic) in the IM140815 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



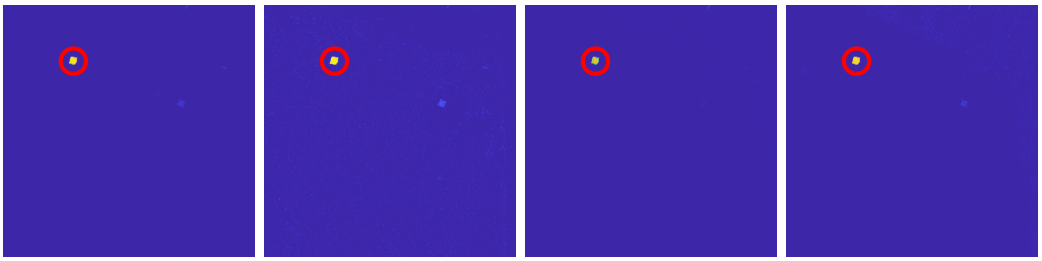
(a) (b) (c) (d)

Figure B.6.45: JSSDR-TD results detecting target 4 (green perspex) in the IM140815 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



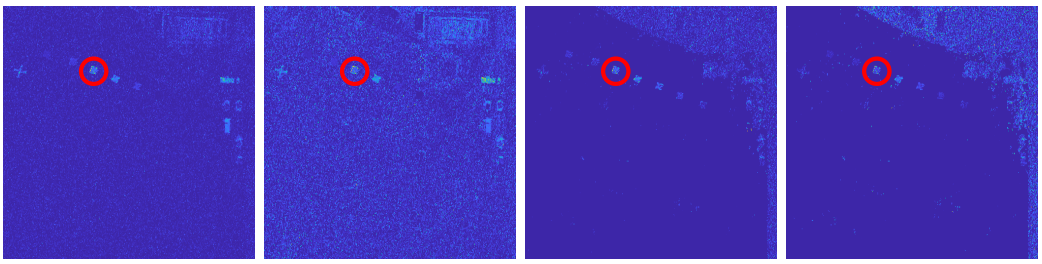
(a) (b) (c) (d)

Figure B.6.46: JSSDR-TD results detecting target 5 (grey ceramic) in the IM140815 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

Figure B.6.47: JSSDR-TD results detecting target 6 (orange perspex) in the IM140813 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.



(a) (b) (c) (d)

Figure B.6.48: JSSDR-TD results detecting target 7 (white perspex) in the IM140813 Image with **a)** Raw ($L = 80$) data, **b)** ICA ($K = 20$), **c)** Raw data and NDVI_{re}, and **d)** ICA and NDVI_{re}.

B.7 Additional CASSI-TD Macbeth Results

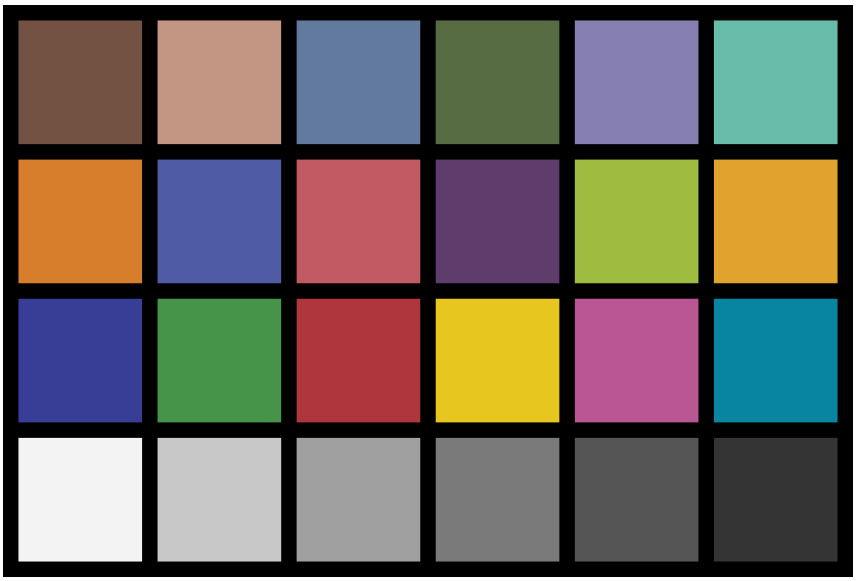


Figure B.7.1: Synthetic Gretag-Macbeth Colour Checker.

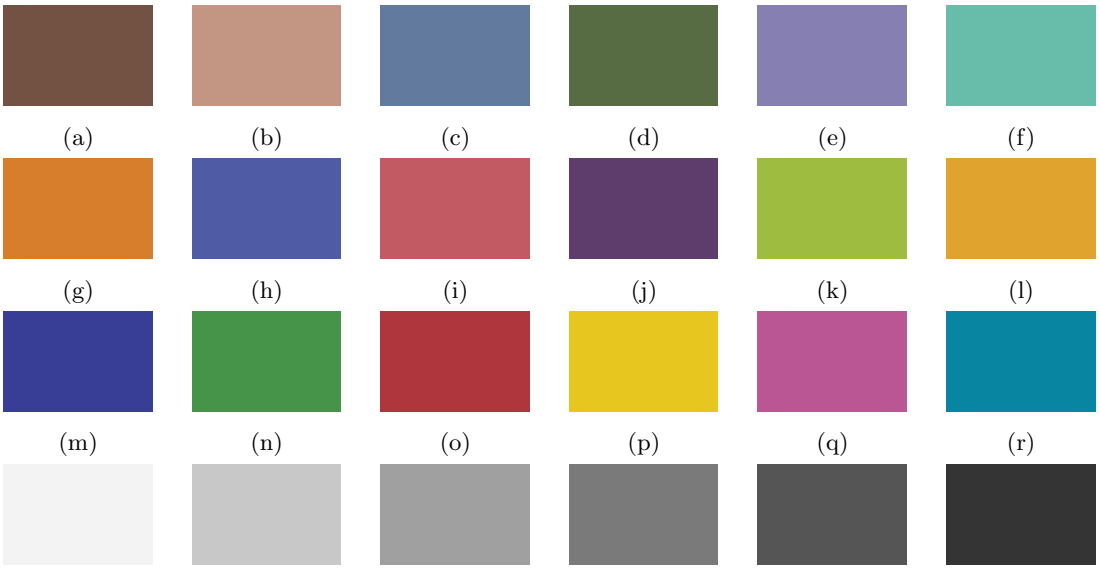


Figure B.7.2: Target images created using each of the spectra from Figure B.7.1.

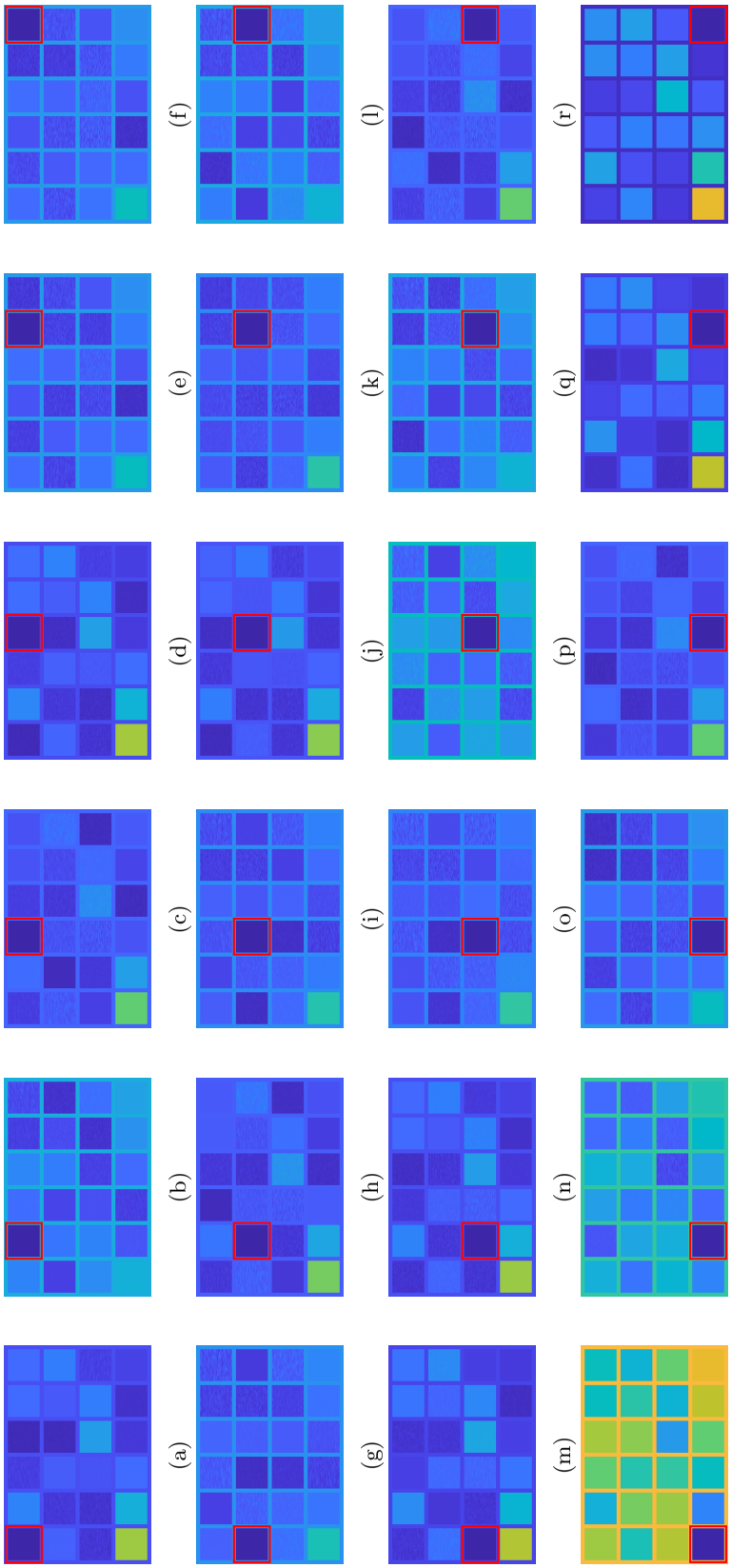


Figure B.7.3: Absolute difference images between Figure B.7.1 and each of the target images in Figure B.7.2 using a Boolean CA and $K = 10$.

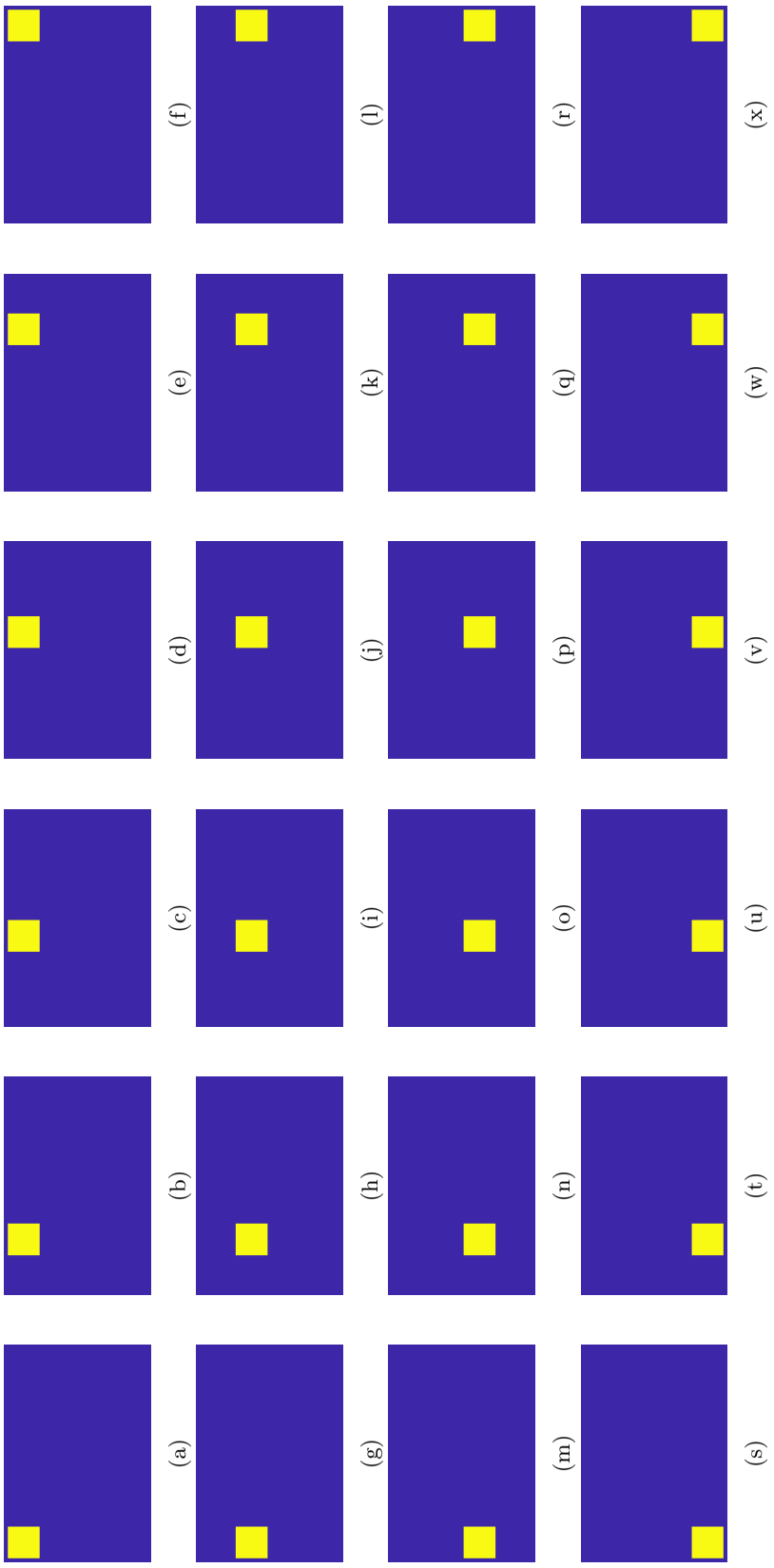
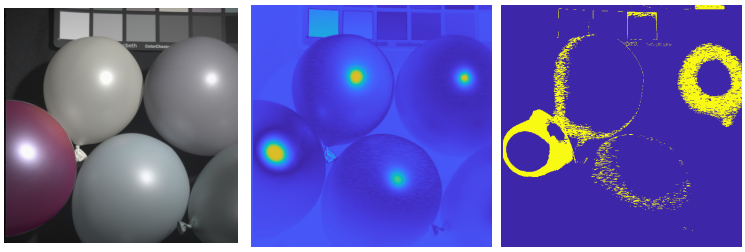


Figure B.7.4: CASSI-TD results on the images from Figure B.7.3 with $T = 0.99$.

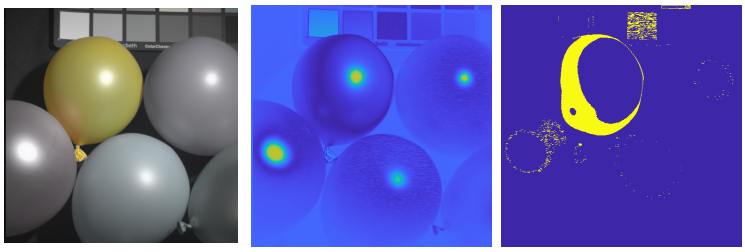
B.8 Additional CASSI-TD CAVE Results

B.8.1 CAVE "balloons"



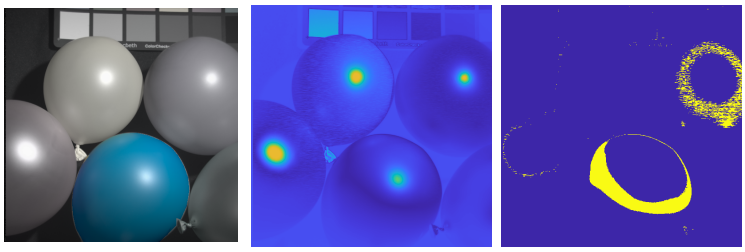
(a) (b) (c)

Figure B.8.1: CASSI-TD result detecting target 1 (pink balloon) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



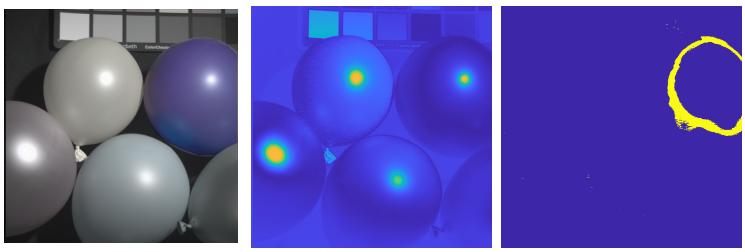
(a) (b) (c)

Figure B.8.2: CASSI-TD result detecting target 2 (yellow balloon) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



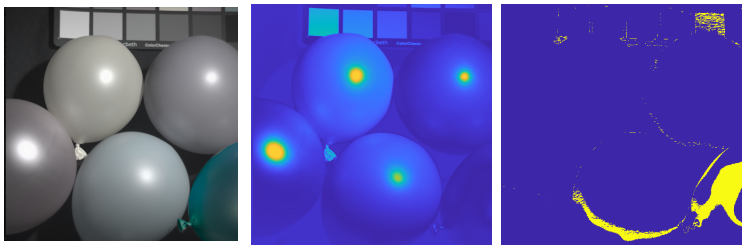
(a) (b) (c)

Figure B.8.3: CASSI-TD result detecting target 3 (cyan balloon) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



(a) (b) (c)

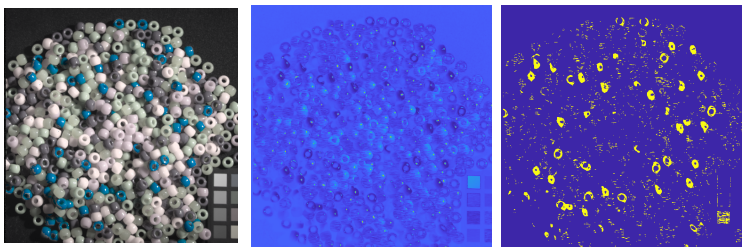
Figure B.8.4: CASSI-TD result detecting target 4 (purple balloon) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.99$.



(a) (b) (c)

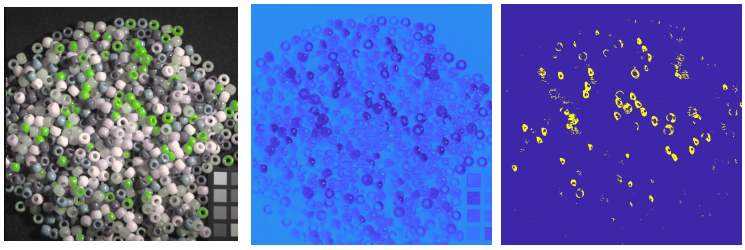
Figure B.8.5: CASSI-TD result detecting target 5 (teal balloon) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.99$.

B.8.2 CAVE "beads"



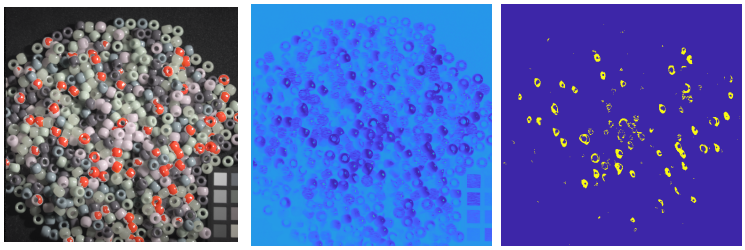
(a) (b) (c)

Figure B.8.6: CASSI-TD result detecting target 1 (blue beads) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.



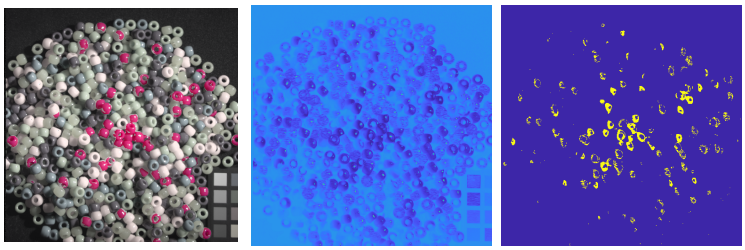
(a) (b) (c)

Figure B.8.7: CASSI-TD result detecting target 2 (green beads) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.



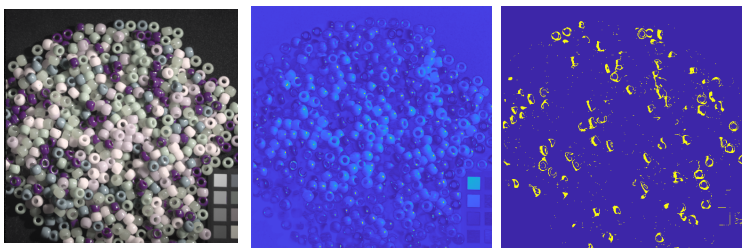
(a) (b) (c)

Figure B.8.8: CASSI-TD result detecting target 3 (orange beads) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.



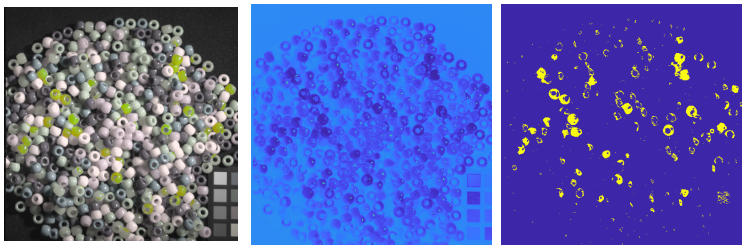
(a) (b) (c)

Figure B.8.9: CASSI-TD result detecting target 4 (pink beads) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.



(a) (b) (c)

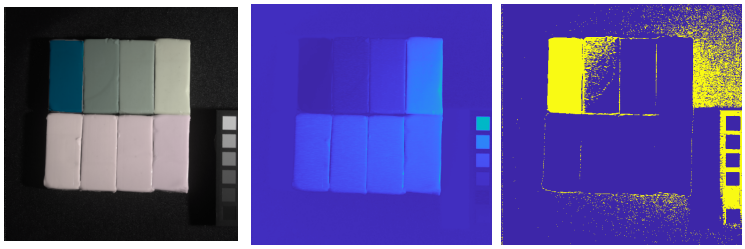
Figure B.8.10: CASSI-TD result detecting target 5 (purple beads) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



(a) (b) (c)

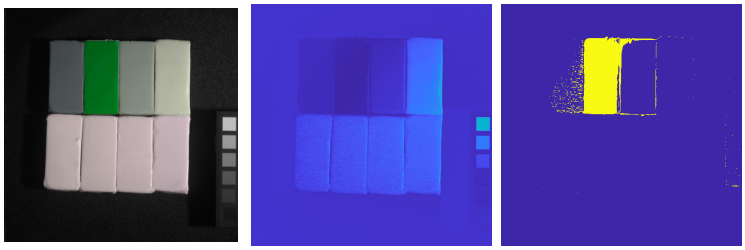
Figure B.8.11: CASSI-TD result detecting target 6 (yellow beads) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.

B.8.3 CAVE "clay"



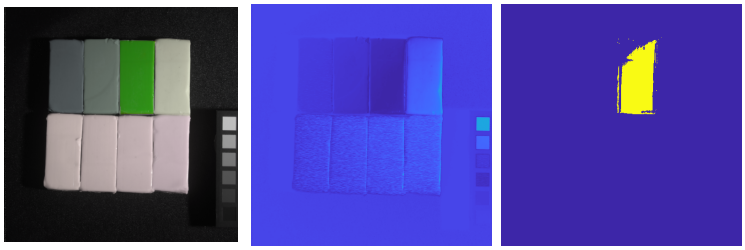
(a) (b) (c)

Figure B.8.12: CASSI-TD result detecting target 1 (blue clay) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



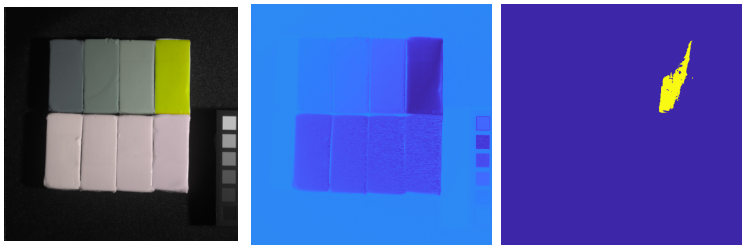
(a) (b) (c)

Figure B.8.13: CASSI-TD result detecting target 2 (green clay) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



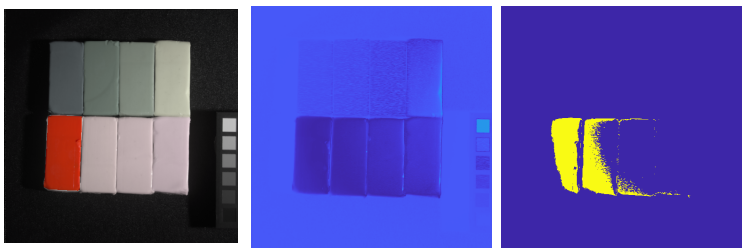
(a) (b) (c)

Figure B.8.14: CASSI-TD result detecting target 3 (lime clay) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



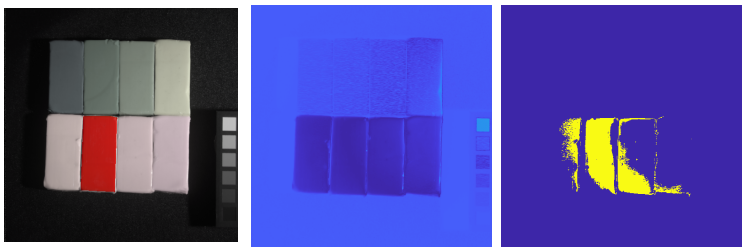
(a) (b) (c)

Figure B.8.15: CASSI-TD result detecting target 4 (yellow clay) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



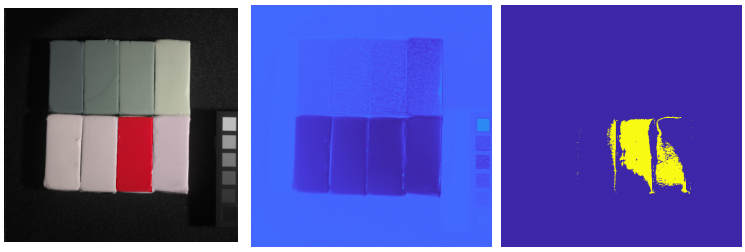
(a) (b) (c)

Figure B.8.16: CASSI-TD result detecting target 5 (orange clay) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



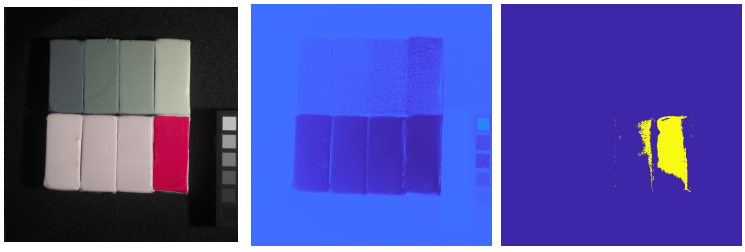
(a) (b) (c)

Figure B.8.17: CASSI-TD result detecting target 6 (scarlet clay) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



(a) (b) (c)

Figure B.8.18: CASSI-TD result detecting target 7 (red clay) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



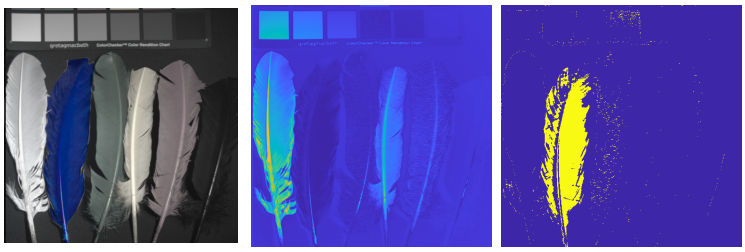
(a) (b) (c)

Figure B.8.19: CASSI-TD result detecting target 8 (pink clay) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

B.8.4 CAVE "feathers"

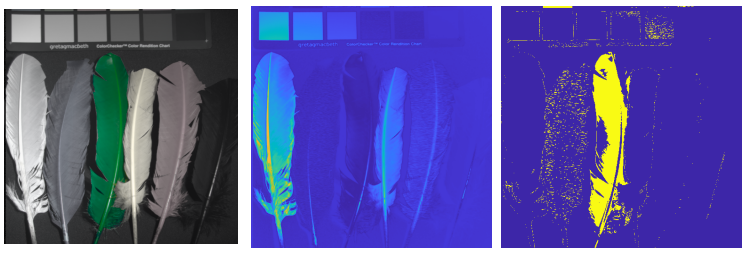
(a) (b) (c)

Figure B.8.20: CASSI-TD result detecting target 1 (white feather) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.



(a) (b) (c)

Figure B.8.21: CASSI-TD result detecting target 2 (blue feather) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.

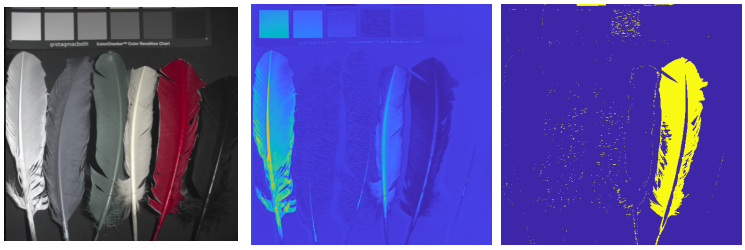


(a) (b) (c)

Figure B.8.22: CASSI-TD result detecting target 3 (green feather) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.

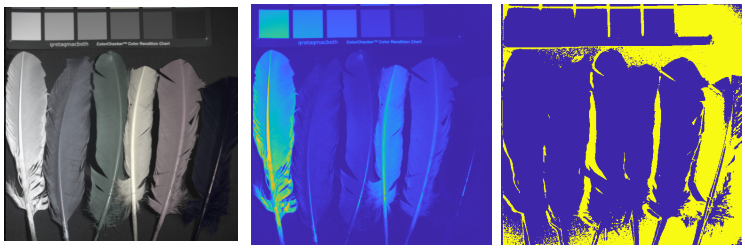
(a) (b) (c)

Figure B.8.23: CASSI-TD result detecting target 4 (yellow feather) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.



(a) (b) (c)

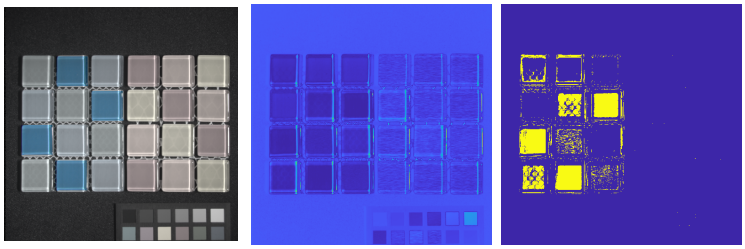
Figure B.8.24: CASSI-TD result detecting target 5 (red feather) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



(a) (b) (c)

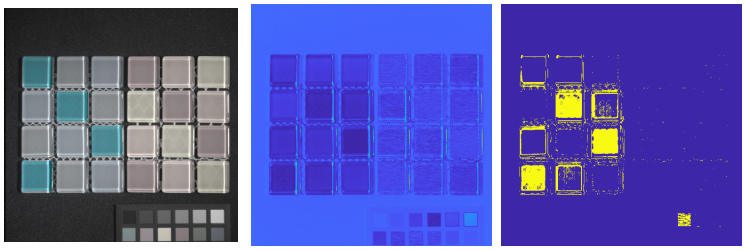
Figure B.8.25: CASSI-TD result detecting target 6 (black feather) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.99$.

B.8.5 CAVE "glass tiles"



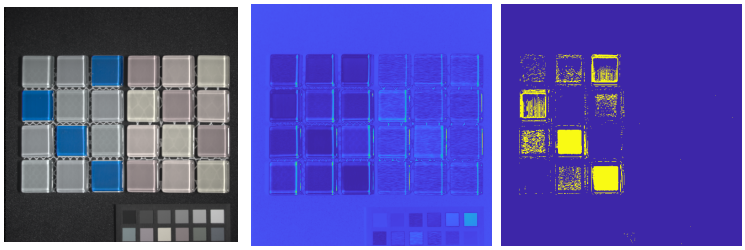
(a) (b) (c)

Figure B.8.26: CASSI-TD result detecting target 1 (teal tiles) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



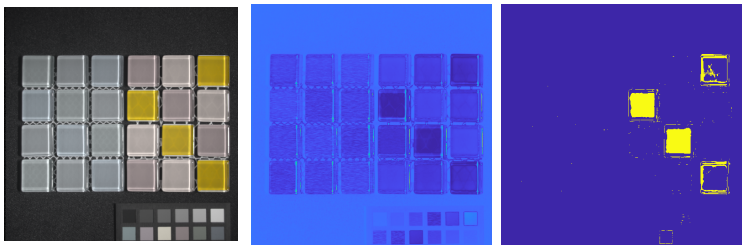
(a) (b) (c)

Figure B.8.27: CASSI-TD result detecting target 2 (cyan tiles) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



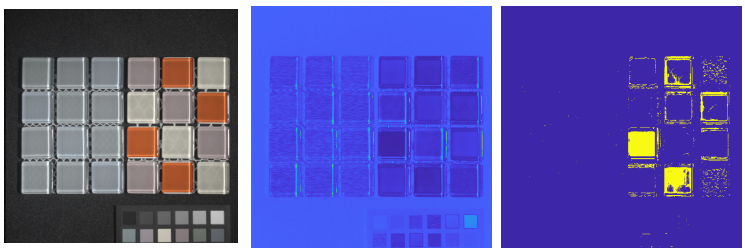
(a) (b) (c)

Figure B.8.28: CASSI-TD result detecting target 3 (blue tiles) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



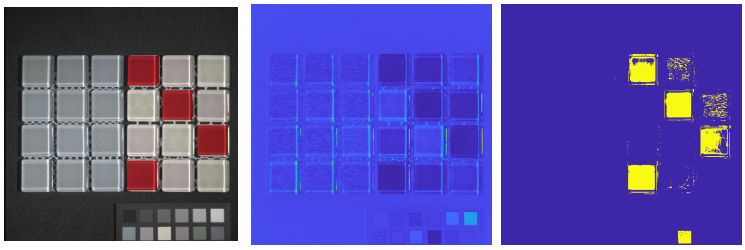
(a) (b) (c)

Figure B.8.29: CASSI-TD result detecting target 4 (yellow tiles) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



(a) (b) (c)

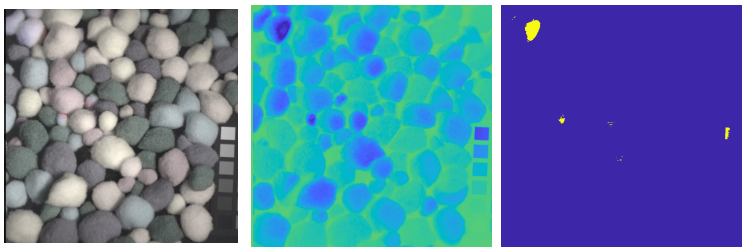
Figure B.8.30: CASSI-TD result detecting target 5 (orange tiles) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



(a) (b) (c)

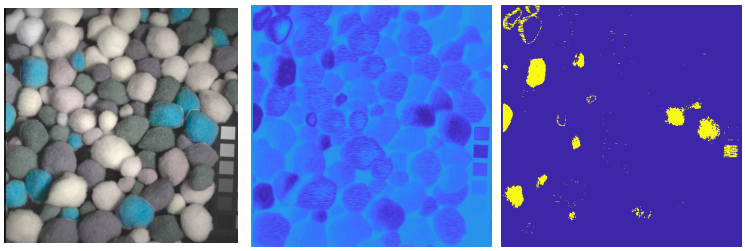
Figure B.8.31: CASSI-TD result detecting target 6 (red tiles) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.

B.8.6 CAVE "pompoms"



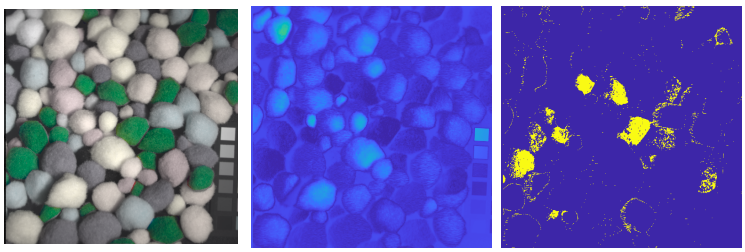
(a) (b) (c)

Figure B.8.32: CASSI-TD result detecting target 1 (white pompoms) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.



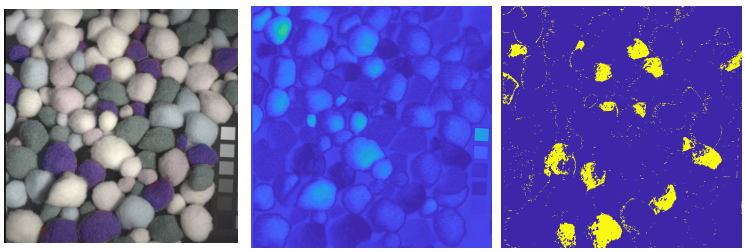
(a) (b) (c)

Figure B.8.33: CASSI-TD result detecting target 2 (blue pompoms) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.



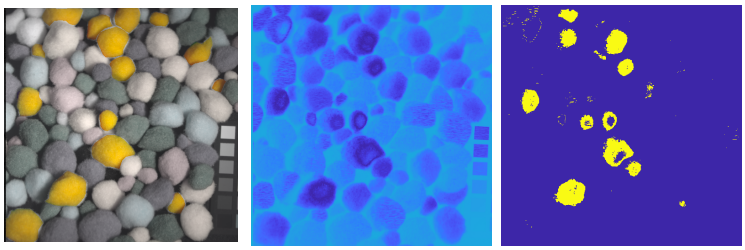
(a) (b) (c)

Figure B.8.34: CASSI-TD result detecting target 3 (green pompoms) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



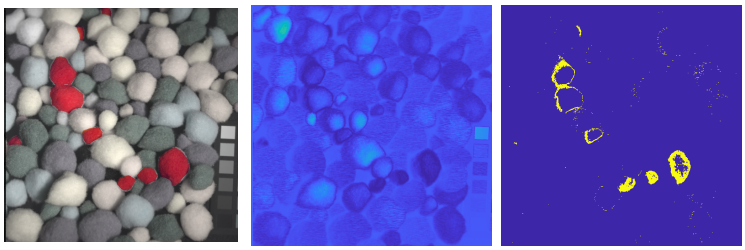
(a) (b) (c)

Figure B.8.35: CASSI-TD result detecting target 4 (purple pompoms) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



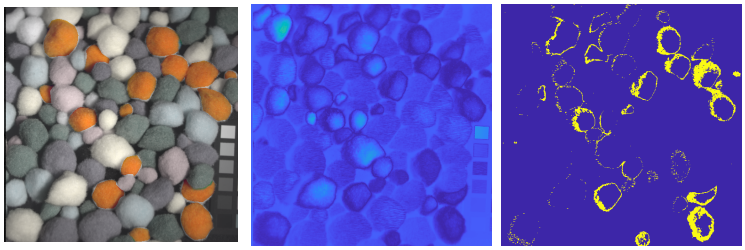
(a) (b) (c)

Figure B.8.36: CASSI-TD result detecting target 5 (yellow pompoms) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.



(a) (b) (c)

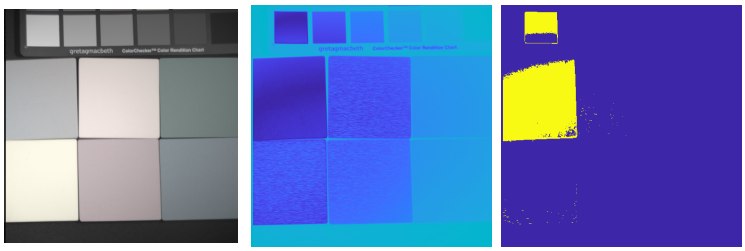
Figure B.8.37: CASSI-TD result detecting target 6 (red pompoms) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



(a) (b) (c)

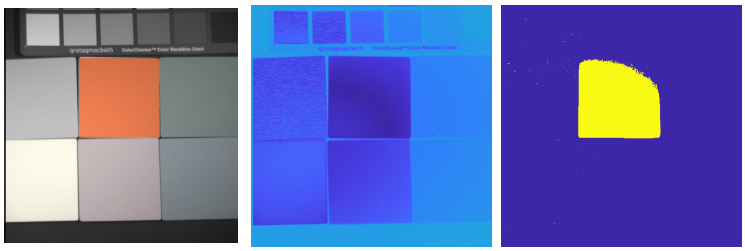
Figure B.8.38: CASSI-TD result detecting target 7 (orange pompoms) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

B.8.7 CAVE "sponges"



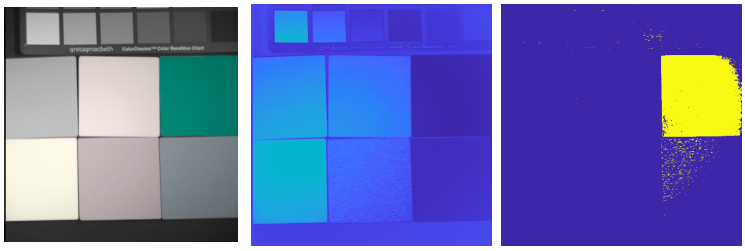
(a) (b) (c)

Figure B.8.39: CASSI-TD result detecting target 1 (grey sponge) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.



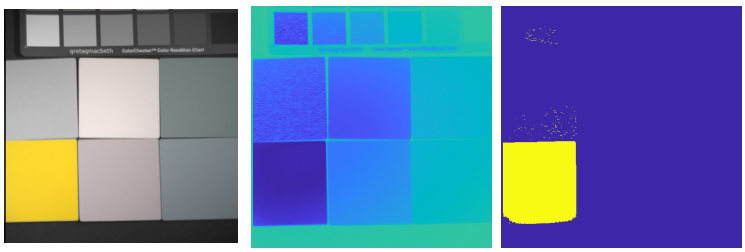
(a) (b) (c)

Figure B.8.40: CASSI-TD result detecting target 2 (orange sponge) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.



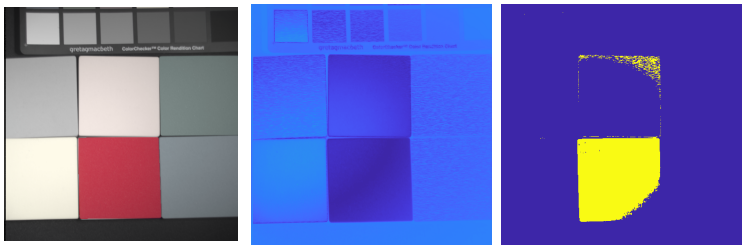
(a) (b) (c)

Figure B.8.41: CASSI-TD result detecting target 3 (green sponge) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



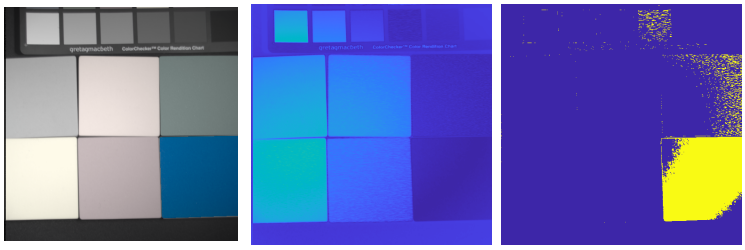
(a) (b) (c)

Figure B.8.42: CASSI-TD result detecting target 4 (yellow sponge) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.



(a) (b) (c)

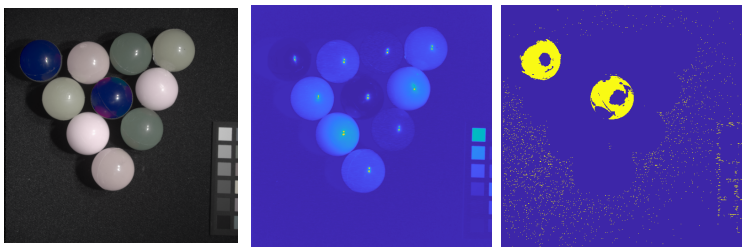
Figure B.8.43: CASSI-TD result detecting target 5 (red sponge) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.



(a) (b) (c)

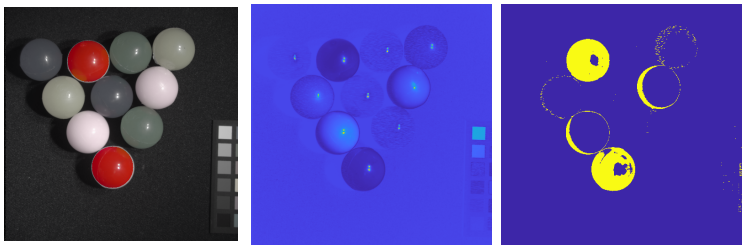
Figure B.8.44: CASSI-TD result detecting target 6 (blue sponge) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

B.8.8 CAVE "superballs"



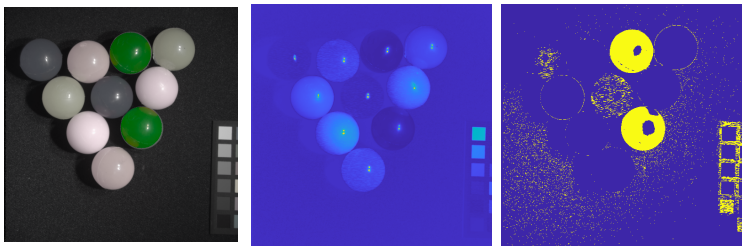
(a) (b) (c)

Figure B.8.45: CASSI-TD result detecting target 1 (blue balls) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.99$.



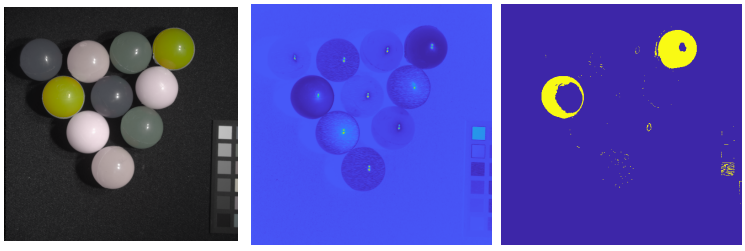
(a) (b) (c)

Figure B.8.46: CASSI-TD result detecting target 2 (red balls) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



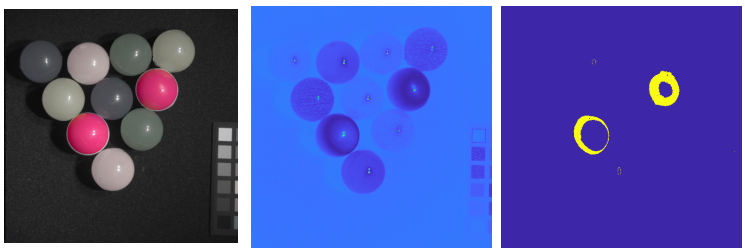
(a) (b) (c)

Figure B.8.47: CASSI-TD result detecting target 3 (green balls) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



(a) (b) (c)

Figure B.8.48: CASSI-TD result detecting target 4 (yellow balls) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

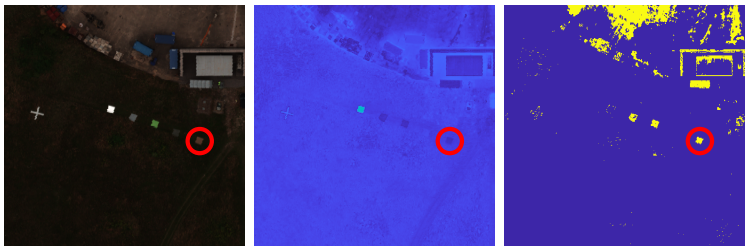


(a) (b) (c)

Figure B.8.49: CASSI-TD result detecting target 5 (pink balls) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and highlighted target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

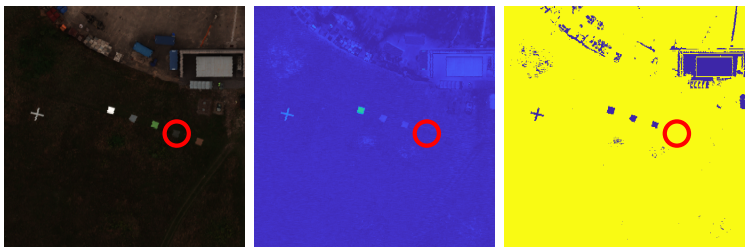
B.9 Additional CASSI-TD UDRC Results

B.9.1 IM140804



(a) (b) (c)

Figure B.9.1: CASSI-TD result detecting target 1 (brown carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.

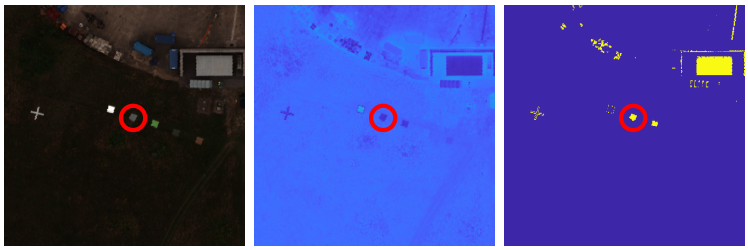


(a) (b) (c)

Figure B.9.2: CASSI-TD result detecting target 2 (green carpet)using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.

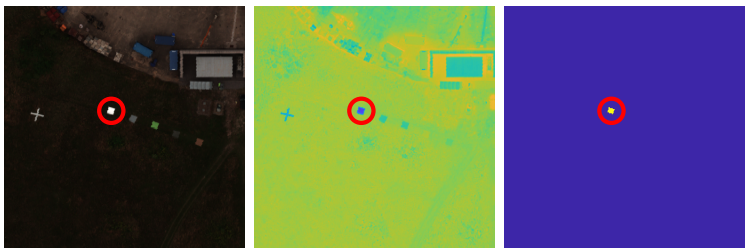
(a) (b) (c)

Figure B.9.3: CASSI-TD result detecting target 3 (green ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.



(a) (b) (c)

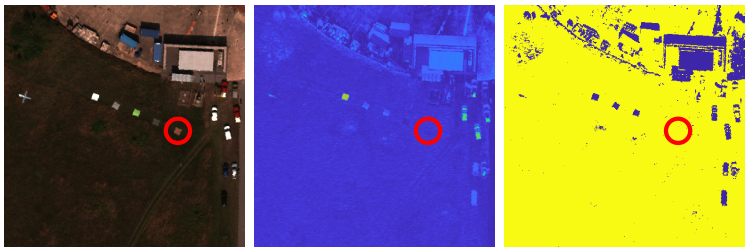
Figure B.9.4: CASSI-TD result detecting target 4 (grey ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.



(a) (b) (c)

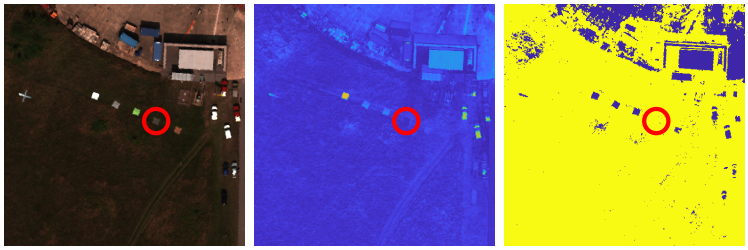
Figure B.9.5: CASSI-TD result detecting target 5 (white perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.

B.9.2 IM140806



(a) (b) (c)

Figure B.9.6: CASSI-TD result detecting target 1 (brown carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.



(a) (b) (c)

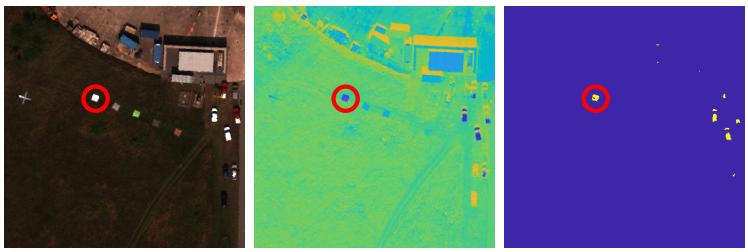
Figure B.9.7: CASSI-TD result detecting target 2 (green carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.

(a) (b) (c)

Figure B.9.8: CASSI-TD result detecting target 3 (green ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.

(a) (b) (c)

Figure B.9.9: CASSI-TD result detecting target 5 (grey ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.



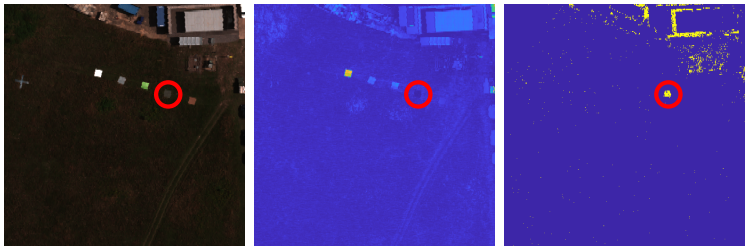
(a) (b) (c)

Figure B.9.10: CASSI-TD result detecting target 7 (grey ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.

B.9.3 IM140807

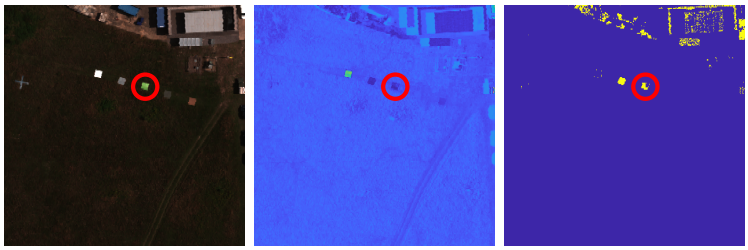
(a) (b) (c)

Figure B.9.11: CASSI-TD result detecting target 1 (brown carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



(a) (b) (c)

Figure B.9.12: CASSI-TD result detecting target 2 (green carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.99$.

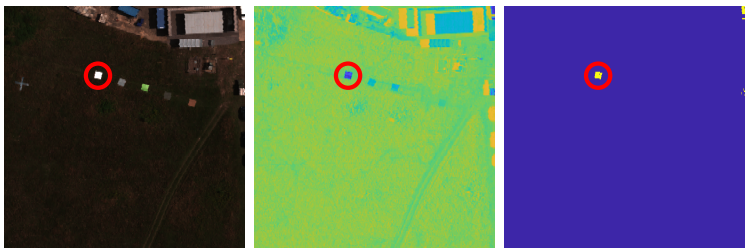


(a) (b) (c)

Figure B.9.13: CASSI-TD result detecting target 3 (green ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

(a) (b) (c)

Figure B.9.14: CASSI-TD result detecting target 5 (grey ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



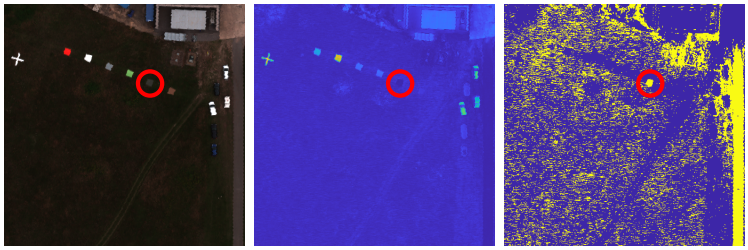
(a) (b) (c)

Figure B.9.15: CASSI-TD result detecting target 7 (white perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.

B.9.4 IM140808

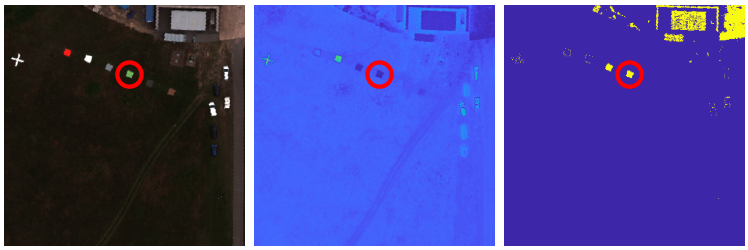
(a) (b) (c)

Figure B.9.16: CASSI-TD result detecting target 1 (brown carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



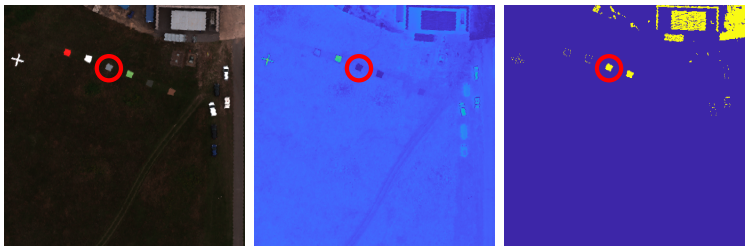
(a) (b) (c)

Figure B.9.17: CASSI-TD result detecting target 2 (green carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



(a) (b) (c)

Figure B.9.18: CASSI-TD result detecting target 3 (green ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.

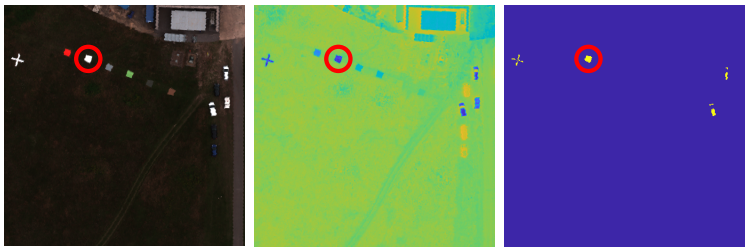


(a) (b) (c)

Figure B.9.19: CASSI-TD result detecting target 5 (grey ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.

(a) (b) (c)

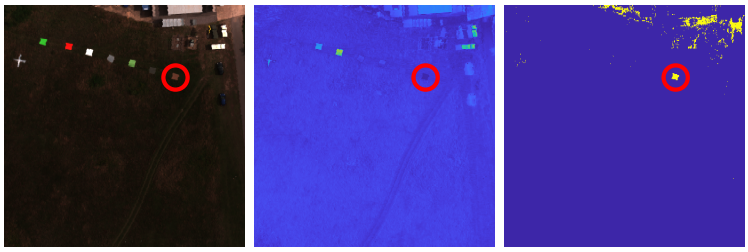
Figure B.9.20: CASSI-TD result detecting target 6 (orange perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.



(a) (b) (c)

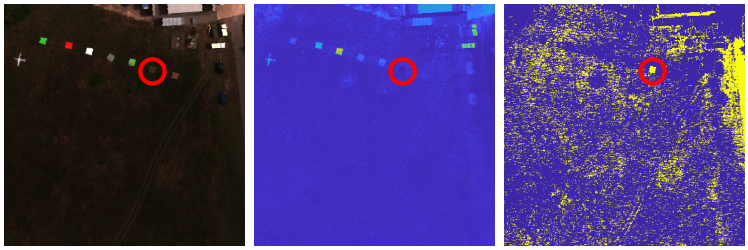
Figure B.9.21: CASSI-TD result detecting target 7 (white perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.85$.

B.9.5 IM140812



(a) (b) (c)

Figure B.9.22: CASSI-TD result detecting target 1 (brown carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.

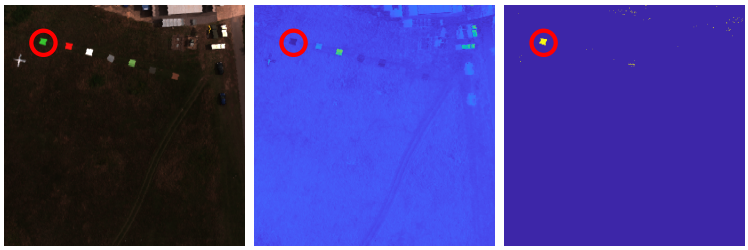


(a) (b) (c)

Figure B.9.23: CASSI-TD result detecting target 2 (green carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.

(a) (b) (c)

Figure B.9.24: CASSI-TD result detecting target 3 (green ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



(a) (b) (c)

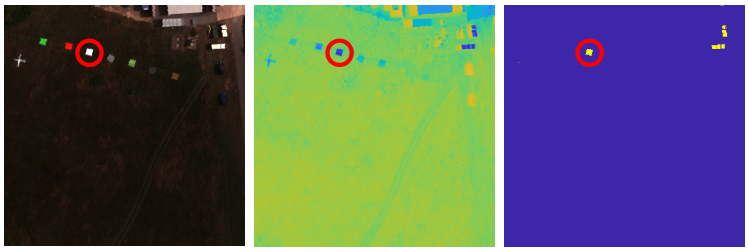
Figure B.9.25: CASSI-TD result detecting target 4 (green perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

(a) (b) (c)

Figure B.9.26: CASSI-TD result detecting target 5 (grey ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

(a) (b) (c)

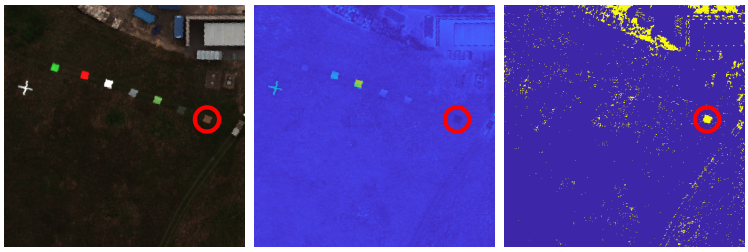
Figure B.9.27: CASSI-TD result detecting target 6 (orange perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.



(a) (b) (c)

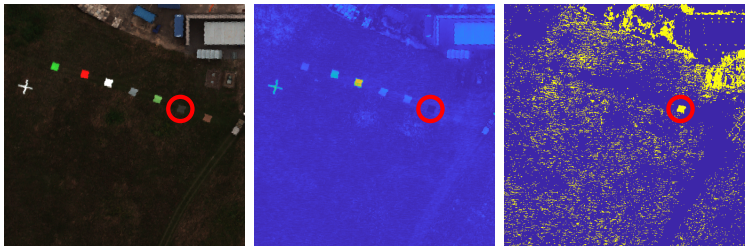
Figure B.9.28: CASSI-TD result detecting target 7 (white perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.

B.9.6 IM140813



(a) (b) (c)

Figure B.9.29: CASSI-TD result detecting target 1 (brown carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

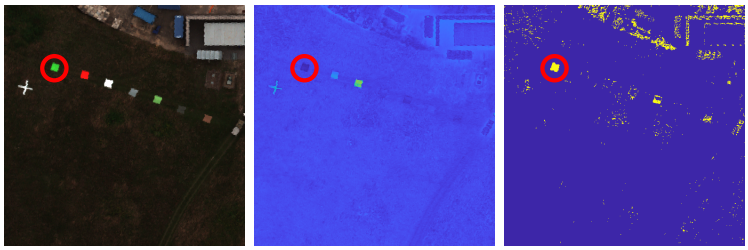


(a) (b) (c)

Figure B.9.30: CASSI-TD result detecting target 2 (green carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.

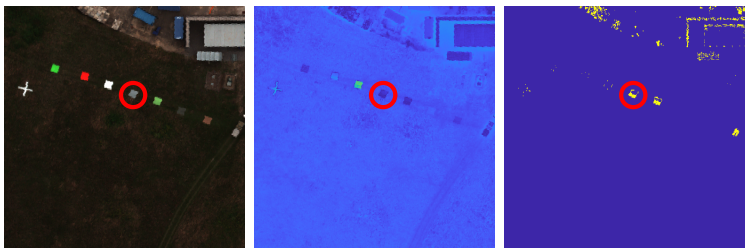
(a) (b) (c)

Figure B.9.31: CASSI-TD result detecting target 3 (green ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



(a) (b) (c)

Figure B.9.32: CASSI-TD result detecting target 4 (green perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.

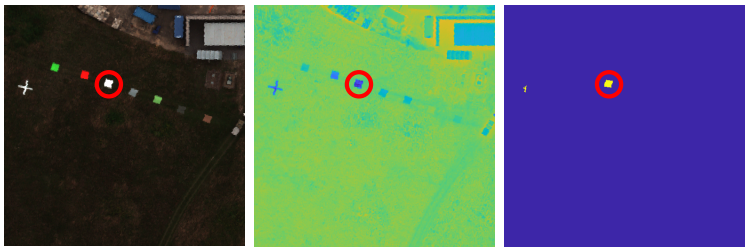


(a) (b) (c)

Figure B.9.33: CASSI-TD result detecting target 5 (grey ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

(a) (b) (c)

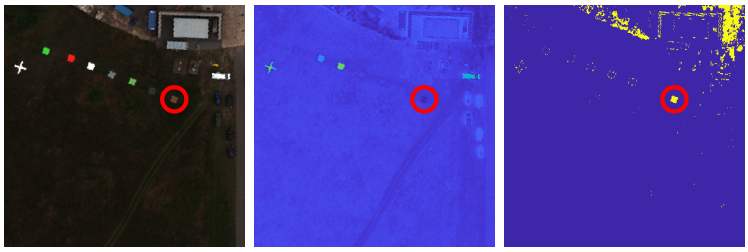
Figure B.9.34: CASSI-TD result detecting target 6 (orange perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.



(a) (b) (c)

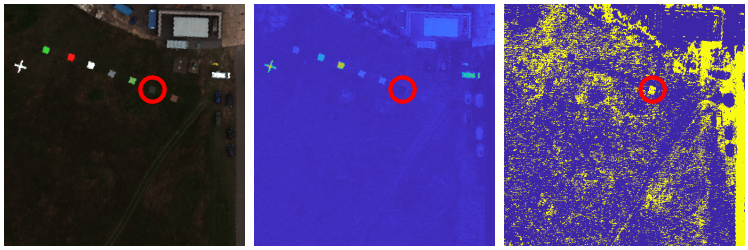
Figure B.9.35: CASSI-TD result detecting target 7 (white perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.8$.

B.9.7 IM140815



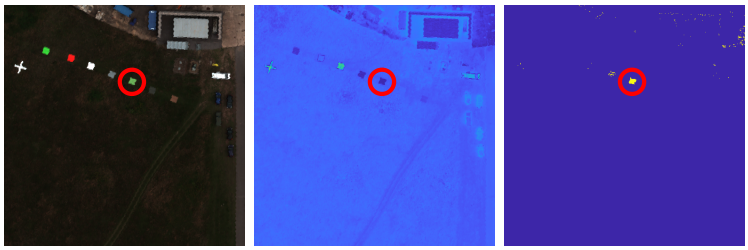
(a) (b) (c)

Figure B.9.36: CASSI-TD result detecting target 1 (brown carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.



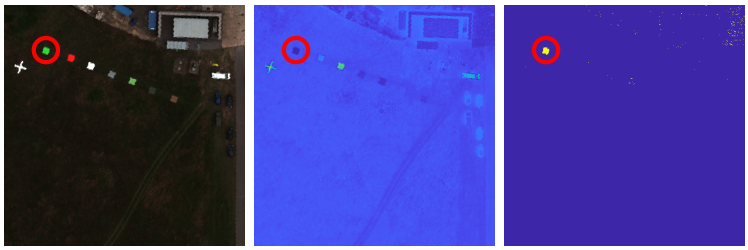
(a) (b) (c)

Figure B.9.37: CASSI-TD result detecting target 2 (green carpet) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



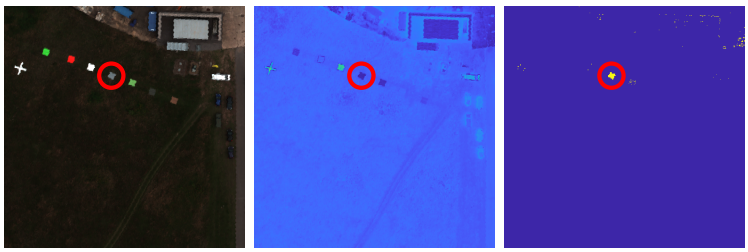
(a) (b) (c)

Figure B.9.38: CASSI-TD result detecting target 3 (green ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.



(a) (b) (c)

Figure B.9.39: CASSI-TD result detecting target 4 (green perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.95$.

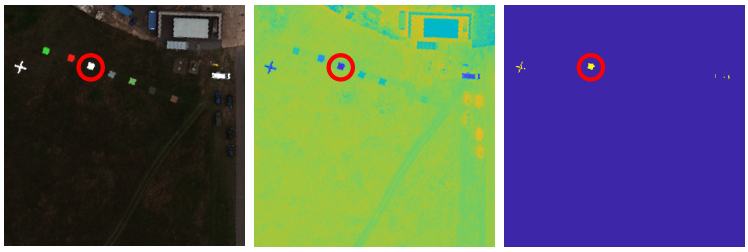


(a) (b) (c)

Figure B.9.40: CASSI-TD result detecting target 5 (grey ceramic) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.97$.

(a) (b) (c)

Figure B.9.41: CASSI-TD result detecting target 6 (orange perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.



(a) (b) (c)

Figure B.9.42: CASSI-TD result detecting target 7 (white perspex) using DD-CASSI with a Boolean CA where $K = 10$. **a)** Image and target. **b)** $\boldsymbol{Y}_{\Delta}$. **c)** Detection with $T = 0.9$.

Appendix C

List of Author's Publications

C.1 Journal publications

C.1.1 **F. Macfarlane**, P. Murray, S. Marshall, B. Perret, A. N. Evans, and H. White, "Robust object detection in colour images using a multivariate percentage occupancy hit-or-miss transform," *Mathematical Morphology - Theory and Applications*, 5 (1). pp. 128-152. ISSN 2353-3390

C.1.2 **F. Macfarlane**, P. Murray, S. Marshall, and H. White, "Investigating the Effects of a Combined Spatial and Spectral Dimensionality Reduction Approach for Aerial Hyperspectral Target Detection Applications" Remote Sensing. 2021; 13(9):1647. https://doi.org/10.3390/rs13091647

C.1.3 S. Song, D. Gibson, S. Ahmadzadeh, H.-O. Chu, D. Hutson, B. Warden, R. Overend, **F. Macfarlane**, P. Murray, S. Marshall, M. Aitkenhead, D. Bienkowski, and R. Allison,"Low cost Hyper-spectral imaging system using Linear Variable Bandpass Filter for Agritech Applications", Appl. Opt. 59, A167-A175 (2020)

C.2 Conference publications

C.2.1 **F. Macfarlane**, P. Murray, S. Marshall, and H. White, "Object detection and classification in aerial hyperspectral imagery using a multivariate hit-or-miss transform", in *SPIE Defense + Commercial Sensing (DCS 2019)*, Baltimore, Maryland, United States, 2019.

C.2.2 **F. Macfarlane**, P. Murray, S. Marshall, and H. White, " A fast hyperspectral hit-or-miss transform with integrated projection-based dimensionality reduction", in *7th Hyperspectral Imaging and Applications Conference (HSI 2018)*, Coventry, United Kingdom, 2018.

C.2.3 **F. Macfarlane**, P. Murray, S. Marshall, B. Perret, A. N. Evans, and H. White, "A Colour Hit-or-Miss Transform Based on a Rank Ordered Distance Measure", in *2018 26th European Signal Processing Conference (EUSIPCO)*, Rome, Italy, 2018.