論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Small-sized object detection via hyperspectral date
著者(和文)	厳露
Author(English)	Lu Yan
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11674号, 授与年月日:2020年12月31日, 学位の種別:課程博士, 審査員:山口 雅浩,熊澤 逸夫,金子 寛彦,永井 岳大,渡辺 義浩
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11674号, Conferred date:2020/12/31, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Small-sized object detection via hyperspectral data

by

Lu Yan

A dissertation submitted to
Tokyo Institute of Technology
for the degree of

DOCTOR OF ENGINEERING

Adviser: Professor Masahiro Yamaguchi

Table of content

Table of content ..... **i**

List of Figures ..... **iv**

List of Tables ..... **viii**

Abstract ..... **ix**

Acknowledgment ..... **xii**

Chapter 1: Introduction **1**

1.1 Research background *1*

1.2 Remaining issues *2*

1.3 Research purpose *3*

1.4 Structure of the thesis *3*

Chapter 2: Object detection on the hyperspectral dataset **5**

2.1 Hyperspectral Dataset *5*

2.1.1 Dimension reduction 7

2.1.2 Spectral Unmixing 8

2.1.3 Endmember extraction 9

2.2 Object detection methods *10*

2.2.1 Supervised methods 11

2.2.2 Unsupervised methods 11

2.2.3 Semi-supervised methods 13

Chapter 3: Using hyperspectral image enhancement method for small size object detection on the sea surface 15

3.1 Research Background 16

3.2 Detection Method 19

3.2.1 Overview of the proposed method 19

3.2.2 Background separation 20

3.2.3 Endmember spectra extraction 21

3.2.4 Classification 22

3.3 Experiment 24

3.3.1 Materials 24

3.3.2 Processing 25

3.4 Results 29

3.5 Discussion 32

Chapter 4: Effect of the restoration of saturated signals in hyperspectral image analysis and color reproduction 33

4.1 Background 33

4.2 Methodology 37

4.2.1 Saturated signals in hyperspectral imaging.....37

4.2.2 Restoration of saturated hyperspectral signals.....39

4.3 Experimental Results 43

4.3.1 Object detection 43

4.3.2 Color reproduction application 50

4.3.3 Simulated saturated hyperspectral data in color reproduction application.....56

4.4 Conclusion 63

Chapter 5: A novel two-stage deep learning-based small-object detection using hyperspectral images 65

5.1 Background..... 65

5.2 Hyperspectral Small-sized object detection..... 69

5.3 Proposed Deep Learning architecture..... 71

5.3.1 Dataset.....75

5.4 Experimental Results and Comparisons 77

5.5 Conclusions..... 83

Chapter 6: Conclusion and future work.....85

References.....89

List of Publications 101

List of Figures

Figure 1.1 The Structure of this thesis 4

Figure 2.1 Hyperspectral Image Data Cube Demo 6

Figure 2.2 Conceptual figure of airborne HIS for human rescue 7

Figure 2.3 Hyperspectral Image Data Cube Correlation between bands.....8

Figure 3.1 Proposed processing chain 20

Figure 3.2 HSI used in this study, sub-region within the white rectangle is the testing area. Human positions have been marked as green rectangles. 26

Figure 3.3 The principal component cumulative contribution.....27

Figure 3.4 HSI images at wavelength. (a).495 nm and (b)695 nm; images after removing first 3 principal components at the wavelength (c) 495 nm and (d) 695nm 28

Figure 3.5 (a) Original Spectral of Objects presented in the scene. (b) Endmember obtained by VCA.....29

Figure 3.6 Results for (a) RX detector, (b) SAM classification, (c) SID classification, left: without removing principal components, right: remove first 3 principal components. The green rectangle is the position of human.....31

Figure 4.1 This figure schematically compares the reflectance calculated from a normal spectrum with the one derived from a saturated signal, which is clipped

at digital value 1000. When both spectra are divided by the reference illumination spectrum, we can observe a hollow in the reflectance derived from the saturated spectrum.....37
Figure 4.2 The flowchart of our method. 43
Figure 4.3 (a) Examples of airborne HIS of the sea surface. Surfboards and humans are observed on the surface. (b) the average spectral data picked from marks 1 – 4 in the image (a).....44
Figure 4.5 Selected examples of spectra. The blue dash-dot line shows the original spectrum, the solid red line the one with the constraint, and the green dash line the one without the constraint. 46
Figure 4.6 Object detection results using the Reed-Xiaoli anomaly detection (RXAD) method. (a) RXAD on the original data and (b) on the restored spectrum. The detected pixels are denoted with red, while the differences in detection are shown in green in (b). 47
Figure 4.7 True positive and false positive comparison for 43 images. The blue circles represent the detection results for the original data, while the red stars the results using the restored data. 48
Figure 4.8 Endmember extracted (a) from the original image and (b) from the restored image 49

Figure 4.9 (a) Saturated versus (b) restored RGB images. Note that the color range in the (a) saturated image was adjusted such that the color of the non-saturated region appears the same in both images. 50

Figure 4.10 (a) Saturated vs. (b) Restored Y channel images(before normalization). Note that the range of color assignment is different. 51

Figure 4.11 Histogram of Y channel of the (a) original saturated and (b) restored images. The horizontal axis shows the luminance value, and the vertical axis indicates the number of pixels in the log scale. 53

Figure 4.12 (a)–(d) Enlarged parts of the original and restored RGB images and the profiles of the RGB values before and after the restoration along the lines indicated by the red arrows. 55

Figure 4.13 RGB images of original data, saturated data, restored data, and the saturated positions (the false color represents the total number of bands saturated in the spectra) where the spectral data were sampled. 57

Figure 4.14. The corresponding spectra of different sub-areas: the original (red), saturated (blue), and restored (green) signals. The labels a-1 ~ 4, b-1 ~ 4, c-1 ~ 4, and d-1 ~ 4 indicate the pixels' positions in Figure 4.13, wherefrom the spectra are obtained. 59

Figure 4.15. CIELAB ΔE of saturated (left) and restored (right) patches. 60

Figure 4.16 Color gamut: the distribution in $L^* - C^*$ plane for each patch corresponding to the patches (A)-(D) in Figure 4.13 62

Figure 4.17 The original (red), saturated (blue), and restored (green) spectra of each rightmost (highest chroma) point shown each patch in Figure 4.16 63

Figure 5.1 Proposed deep neural network structure.....75

Figure 5.2 (a)(b)(c) is sample image from our dataset, (d) is ground truth of (c) with yellow represent the surfboard and the light blue represent the human, the dark blue is the background sea surface 77

Figure 5.3 Spectral Signature sun glint, human, surfboard and sea surface without sun glint 78

Figure 5.4 Detection result on one image 81

Figure 5.5 ROC curve (a) for the demonstrate image (b) leave one out cross validation on 36 images 82

Figure 5.6 Details on human detection performance. Ground truth is shown in column (1), as yellow represent the surfboard, dark blue represents the sea and light blue represent the human. 83

List of Tables

Table 3-1 Detection results (a) Without background constrain, (b) With background constrain 30

Abstract

Hyperspectral images (HSIs) have the advantage of containing spectral and spatial information simultaneously. For various applications, it is important to properly understand the content of HSIs. Many object classification and detection methods have been proposed. Small target detection is considered one of the difficulties. However, there is little work on small-sized object detection and classification using hyperspectral data.

In this thesis, we begin with a brief introduction to HSI, followed by the research background and purpose. This is followed by an overview of related conventional methods in Chapter 2. The rest of the thesis is separated into three chapters.

First, we propose a semi-supervised method for human detection on the sea surface, and the results are discussed in Chapter 3. Small-sized object detection in a vast ocean plays an important role in rescue after accidents or disasters. However, an interested target might occupy only several pixels or less in the image, and due to the limitation of the sensor's resolution, it is difficult to detect the small object. Moreover, the sun glint of the sea surface makes it

even more difficult. In this chapter, we propose an image analysis technique suitable for the computer-aided detection of small objects on the sea surface, especially humans. We first separate objects from the background by adapting a previously proposed image enhancement method and then apply a linear unmixing method to define the endmember's spectrum. At last, we use the spectral angle mapping method to classify presented objects and thus detect small-sized objects. The proposed system provides the following results for supporting the detection of humans and other small objects on the sea surface, an image with spectral color enhancement, alerts of various objects, and human detection results. This multilayered approach is expected to reduce oversight, that is, false negative error. The results of the proposed technique have been compared with existing methods, and our method has successfully enhanced the hyperspectral image. Our method detects small objects from the sea surface with a high human detection rate and shows the ability to further detect humans in this study. The result is less influenced by sun glint effects. The study helps to recognize small objects on the sea surface, and it leads to advances in the rescue system using aircraft equipped HSI technology.

Second, we discovered that the effect of saturated pixels will have a bad influence on the detection method's performance. Thus, we propose a spectral

restoration method to improve the detection of results. In hyperspectral imaging, the captured signal is often affected by saturation due to specular reflection or a peaky spectrum. In this chapter, we propose a restoration method for saturated hyperspectral signals. Our algorithm is based on principal component analysis to obtain the reconstruction basis and then solve a linear constrained least square problem to calculate the coefficients of each basis. We discuss the problems that saturated signals might cause and apply our method to two sets of real hyperspectral images and a set of hyperspectral images with simulated saturation. The results show that our method helps increase unsupervised object detection and improves high-fidelity color reproduction.

Finally, with the advantage of deep learning methods, we propose a novel two-stage deep learning-based hyperspectral small-sized object detection method in Chapter 5. Pixel-wise spectral information is used in the first stage to obtain first-stage classification results, and then the results are combined with spatial information to help eliminate unlikely regions, thereby improving detection accuracy. The method is tested on a dataset of real-world airborne hyperspectral images, and its performance is compared with those of several conventional methods. The results show that the proposed method outperforms current state-of-the-art methods.

Acknowledgment

First of all, I would like to sincerely thank my academic supervisor, Professor Yamaguchi Masahiro, for his continuous support, guidance, and valuable advice through my Ph.D. study. His patience, enthusiasm, and immense knowledge always inspire me and help me in all these years of research.

I also thank all current and former members of the Yamaguchi lab, for their help both in research and daily life.

Last but not least, I would like to thank my family for their support and for giving me time to concentrate on this research.

I would like to express my sincere gratitude to all the people who helped with this research and thesis.

Introduction

1.1 Research background

Remote sensing (RS) is the technology that detects and monitors the physical characteristics of a target or target area without direct contact. RS usually measures the reflected or emitted electromagnetic radiation of the target.

Target detection using remote sensing images plays an important role in remote sensing image processing. The processing procedures consist of acquisition, processing, analysis, and interpretation. It has a wide range of applications, both in civilian and military aspects. In civilian applications, remote sensing technology is widely applied to resource surveys and exploration, environmental pollution monitoring, disaster monitoring, land use planning, crop estimation, meteorological forecasting, urban planning, etc.

According to spectral resolution, RS can be classified as panchromatic (a single band), multispectral (10's of bands), hyperspectral (100's of bands), and ultraspectral (1000's of bands).As the total amount of light incident into the sensor

is fixed, there is always a trade-off between spectral resolution and spatial resolution. Panchromatic images usually have a high spatial resolution, while hyperspectral images have a high spectral resolution but limited spatial resolution.

With the development of the hyperspectral industry, the cost of a hyperspectral sensor has decreased; thus, more civil applications could be carried out nowadays. Portable hyperspectral cameras and unmanned aerial vehicles equipped with hyperspectral cameras could be used in many ways.

Currently, much research has been conducted on hyperspectral image classification, but object detection remains a difficult task, especially small-sized object detection.

1.2 Remaining issues

Given the limited spatial resolution, how to accurately find small-sized targets from the background still has many issues remaining to be solved. These include problems with limit labeling data and how to perform object detection. In addition, little work has been conducted on the influence of saturated pixels: will the

restored spectral improve the detection method's performance remain unanswered. In the conditions of sufficient training data, an accurate detection method is needed. In this thesis, we are trying to solve these issues one by one.

1.3 Research purpose

Hyperspectral images contain rich spectral information that can help distinguish objects. However, the low spatial resolution will prevent the accurate location of the small object. In this study, we will focus on small-sized object detection methods using hyperspectral images. First, we proposed a semi-supervised method that can reduce the human interpretation of the training data; however, the accuracy is not ideal. Detection accuracy is particularly important in rescuing tasks; therefore, we proposed a deep learning-based two-stage detection neural network. This method has greatly improved the detection rate.

1.4 Structure of the thesis

This thesis is organized into six chapters. Chapter 1 introduces the research background, research purpose, and outline of the thesis. Chapter 2 reviews the conventional methods for object detection on hyperspectral images. Chapter 3 describes a proposed semi-supervised method for human detection on the sea

surface. Chapter 4 focuses on the effect of saturated pixels and proposes a spectral restoration method. Results show that this method could help to improve the anomaly detector's performance. Chapter 5 explains a two-stage deep learning-based hyperspectral small-sized object by combining spatial and spectral information of hyperspectral images. Chapter 6 summarizes the proposed methods and concludes them. The structure and relationship are shown in Figure 1.1.

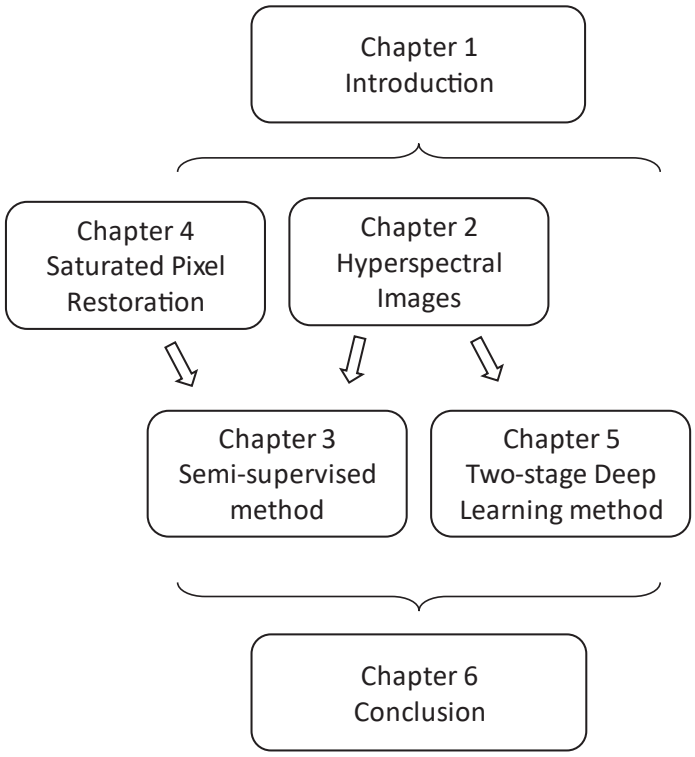


Figure 1.1 The Structure of this thesis

Object detection on the hyperspectral dataset

In this chapter, we will introduce hyperspectral data, which makes up a hyperspectral imaging system (HIS), and explain important notions, such as dimension reduction, spectral unmixing, and endmember extraction. We also introduce several conventional object detection methods using the hyperspectral dataset.

2.1 Hyperspectral dataset

Hyperspectral imaging collects the spectrum for each pixel in a scene. The hyperspectral data cube, as shown in Figure 2.1, along the spectral direction at each wavelength, is a 2D image that contains special information. One pixel has its spectrum, which contains the spectral information. Together, they form a 3D data cube.

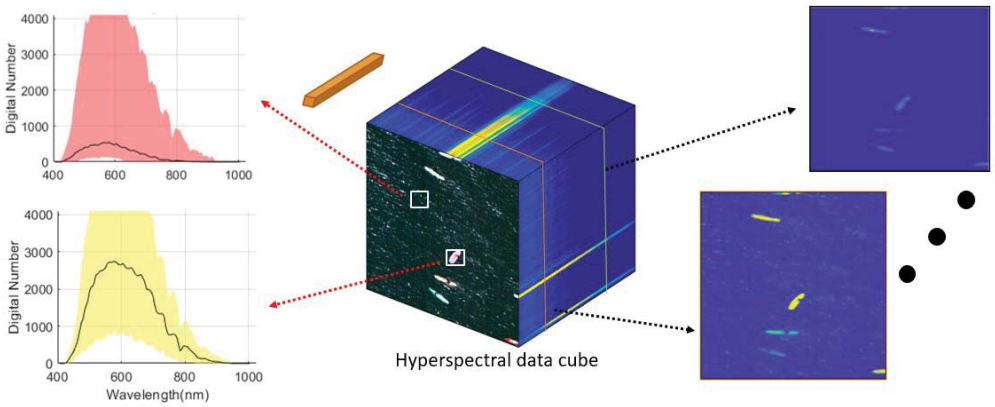


Figure 2.1 Hyperspectral Image Data Cube

Hyperspectral Imaging System (HIS) refers to the system using hyperspectral images as input and provides certain functions. For example, Figure 2.2 illustrates an airborne HIS for human rescue: when the aircraft flies along the sea surface, the HIS processes the real-time imaging collected by hyperspectral

sensors, and once they detect human objects, they will send an alarm to users.

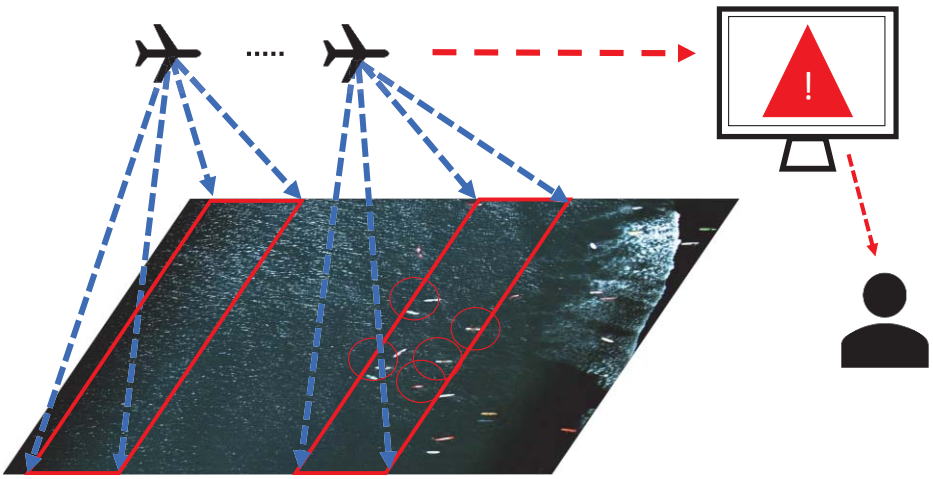


Figure 2.2 Conceptual figure of airborne HIS for human rescue

2.1.1 Dimension reduction

Since hyperspectral images contain hundreds of continuous bands' information, the neighbor wavelengths are usually highly correlated, as shown in Figure 2.3.

There are many dimension reduction methods. Among them, principal component analysis (PCA) is the most well-known and has often been applied in research. PCA is an orthogonal linear transformation and can transform the data to a new coordinate system. Such a coordinate system has the following features: the

data's projection has the greatest variance on the first coordinate, the second greatest variance on the second coordinate, and so on.

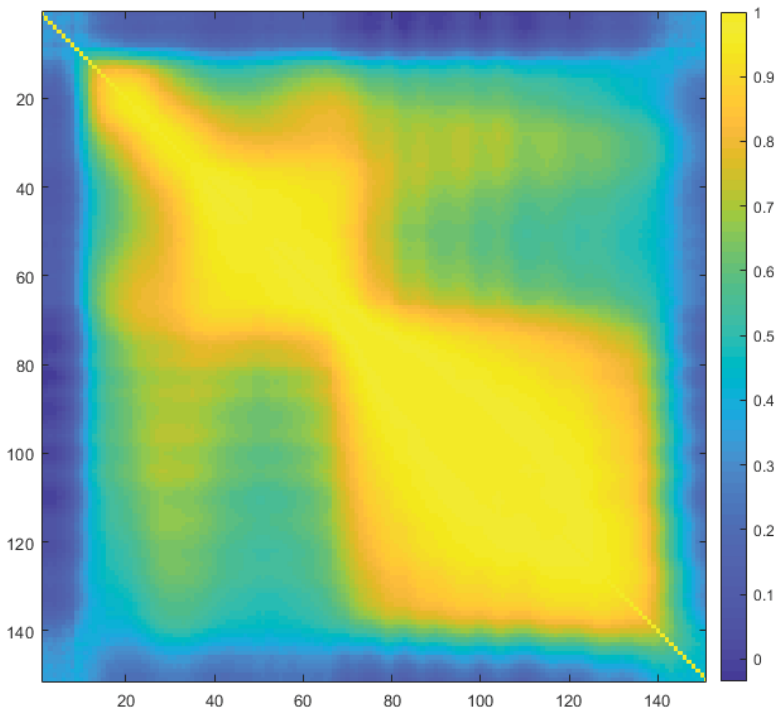


Figure 2.3 Hyperspectral Image Data Cube Correlation between bands

2.1.2 Spectral Unmixing

Hyperspectral unmixing methods can help us define the pure signal of endmembers (the pure signal of one object). There are many hyperspectral unmixing methods. Based on the different hypotheses, there are mainly two kinds of spectral

mixing methods: linear and non-linear. The linear mixing hypothesis supposes that the mixed spectral is composed of a linear combination of several spectral, while the non-linear one believes that the light also interacts within the material and thus cannot be simply presented by linear combinations. Research shows that the non-linear model has the advantage of being more complicated and sophisticated; however, it depends on the model proposed and the parameter selections, while the linear model is easier for calculation and at most of the time performance as well as the non-linear model. In most of the cases, such approximation can be accepted. Hence, here in our study, we will consider the linear unmixing methods.

2.1.3 Endmember extraction

Linear spectral unmixing methods consider that a mixed pixel is the linear combination of endmember signatures. Many algorithms have been proposed for spectral unmixing. Most of the linear spectral unmixing models assume that endmembers are presented in the scene. In the HSI, at least one pure pixel for each endmember exists. Vertex component analysis (VCA) is based on convex sets. It finds the vertices of the simplex after an affine transformation. The obtained vertices are the endmembers. Compared with other linear spectral unmixing methods, VCA has simpler or even better performance, while at the same time its computational

complexity is the lowest.

2.2 Object detection methods

Based on the need or lack of it for training data, detection methods can be classified into three types: supervised, unsupervised, and semi-supervised. Supervised methods provide the system with prior information and then use such information to build classifiers. This kind of method usually performs well with specific requirements: the training data should be well chosen and have a large amount of them. However, when there is an unknown object presented in the detection images, this classifier is mostly unable to work well to define them.

Unsupervised methods are usually applied when there is a lack of training data. In the hyperspectral image detection case, the spectral of objects various influenced by environ light. The status of the sensor can influence spectral signals: it is very difficult to obtain the proper training dataset. Thus, unsupervised methods have the advantage of adapting to different light situations

2.2.1 Supervised methods

Supervised methods require training data or at least some prior knowledge about the spectral characteristics. By learning the features or distribution of the data, we can separate the target from the background. Hybrid sparsity and statistics detectors build a background dictionary with training samples, using a sparse representation-based binary hypothesis model and a generalized likelihood ratio test (GLRT) to obtain a robust detection decision [1]. Even with a priori target spectral signature, the results are greatly affected by the training spectral caused by the spectral variability and are anisotropic in hyperspectral image data. A sparse representation method has been proposed to solve this problem Spectral matching is also performing well in hyperspectral object detection. A target or background spectrum dictionary must be formed. Park et al. [3] used spectral matching to detect ships from the sea surface and performed tests with different similarity measurements.

2.2.2 Unsupervised methods

An unsupervised method for small-sized object detection, also known as anomaly detection (AD), does not need any training data. For hyperspectral imagery,

RXD is known as the benchmark AD [4]. RXD is based on GLRT and under the hypothesis that the background can be modeled as a zero-mean normal distribution and the targets present with low probability. However, this method suffers from high computational cost. When the background is heterogeneous, it may result in a high false alarm rate. To improve the performance and overcome the drawbacks, many improved approaches have been proposed, such as normalized RXD, modified RXD, adaptive subspace detector, etc. [5]. As most statistics-based AD is under the assumption of the linearity of the background while the background could be more complex, nonlinear ADs, such as Kernel RX, were proposed. Kernel-based high-order statistical tensor-based algorithms use the skewness of the image and take not only the distance but also the angle between the abnormal pixel and background pixels [6]. Projection-based AD, such as principal component analysis (PCA), independent component analysis (ICA)-based method, and orthogonal subspace projection (OSP), shows better performance than RXD, with less computational cost [5][7][8]. Geng et al. uses co-skewness tensor to perform higher-order singular value decomposition (SVD) to obtain the projection of bands that represent the target [9]. The wavelet decomposition method can also be applied, such as using a three-dimensional wavelet packet transform [10] or a multidimensional wavelet packet transform with a multiway Wiener filter [11]. The local summation anomaly detection (LSAD) method has been proposed to integrate spectral-spatial

information from neighboring local windows surrounding the pixel under test [12]. Most of the methods are based on relatively large-sized anomalies; however, some are based on small-sized, even subpixel, anomalies. For example, an adaptive Bayesian classifier has been proposed to detect subpixel anomalies. They formed a probabilistic background model, and then the most unlikely pixel was considered to be the target [13]. A graph-based model has been used and classifies a sub-pixel object as boundary pixels, linear pixels, and small encapsulated pixels [14]. A comparison for the quantification of the performance of subpixel object detection has been proposed [15]. To reduce the high computing cost of the AD methods, an approximate computing method for onboard AD has been proposed based on spatial and spectrally down-sampled images [16]. Chang et al. [17] reviewed some unsupervised spectral target analysis and pointed out that when the training data is not reliable, the unsupervised method might be a good choice.

2.2.3 Semi-supervised methods

Semi-supervised methods are those that need limited priority knowledge of the target or background as training data. However, they do not need as much training data as the supervised method. Wang et al. used the spectral projection

method to extract endmembers. They then manually selected the target spectral data. Finally, they applied the variance minimum algorithm to distinguish the target from the background [18]. Yan et al. combined data projection with a spectral unmixing method to detect the small object, the user's need to define the target spectrum. After that, they applied spectral matching methods for the final detection [19].

Method 1: semi-supervised method

Based on Yan, L., et al.: *"Using hyperspectral image enhancement method for small size object detection on the sea surface."* DOI: 10.1117/12.2194606

In rescues after accident or disaster, small size object detection in a vast ocean plays an important role. A hyperspectral imaging system (HIS) is one of the promising approaches. However, the interested target might occupy only several pixels or less in the image due to the limitation of HIS sensor's resolution, it's difficult to detect the small object. Additionally, the sun glint of the sea surface makes the task even harder. In this paper, an image analysis technique suitable for the computer-aided detection of small objects on the sea surface, especially humans has been proposed. Firstly separate objects from the background by adapting a previously proposed image enhancement method. Then a linear unmixing method will be applied to define the endmember's spectrum. At last, the spectral angle mapping method is used to classify presented objects and thus detect small size objects. The proposed system provides the following results to support the detection of humans and other small objects on the sea surface:1) an image with spectral color enhancement, 2) alerts of various objects, and 3) the human detection results. This

multilayered approach is expected to reduce the oversight, i.e., false negative error. We compared the results of the proposed technique with existent methods, and our method has successfully enhanced the hyperspectral image and a higher detection rate of detecting a small object from the sea surface. This shows the ability to further detection of humans in this study). Our method is less influenced by sun glint effects. This study helps to recognize small objects on the sea surface, and it leads to advances in the rescuing system using aircraft equipped HIS technology.

3.1 Research Background

Hyperspectral images (HSIs) have the advantage of providing both spatial and spectral information in high-resolution compared to traditional color images. It has been used in various imaging application systems,. Based on such advantages, hyperspectral images have been increasingly used in various fields such as agriculture, oceanology, geology, etc. Among different applications, object detection and classification have always been an important task of HSI. Normally, object detection algorithms propose to construct a detector based on assumption that the data set follows a prior known probability distribution. There exist several public reflectance spectrum datasets that could be used freely for scientific researches, such as USGS Digital Spectral Library, ASTER Spectral Library, etc. Though, those

databases only cover a minority part of the existent object. For real applications, the prior knowledge of the site presented object are difficult to obtain. Furthermore, in most cases of HSI, the object is very tiny compared to the background size, thus it makes the problem harder to solve.

Human search and rescue missions have always been an important and difficult issue because of its immediacy and accuracy. Sumimoto et al.[20] proposed a method using color information and the composite image sensor to detect the human object in the sea area. Their methods need pre-knowledge of the human object like color, size, or shape. Nevertheless, in most cases, such information cannot be provided properly. Westall et al. [21] have studied several different point target detections techniques and tracking techniques for an aerial search of humans applying different color spaces images. Their results show the possibility to use the machine in the maritime human search application. On the other hand, the system is not yet suitable for real search operations due to the false alarm for an acceptable missed detection rate. Doherty and Rudol [22] propose an Unmanned Aerial Vehicle search and rescue scenario for human body detection. In their proposed method, they used the thermal image to find potential human body temperature alike region. Those regions which not resemble human shape will be rejected. The corresponding pixel in the color image will be identified and using a classifier which been trained with

Haar-like features. Their results show the feasibility of such system; but this kind of system need large dataset of training data and no evidence shows can be used for other rescue situation such as marine. The existing methods need pre-knowledge of human shape or other information, the detection results could be various if the training dataset is not well selected or not enough amount.

The application of HSI is promising as mentioned above, however, several issues need to be solved before having a practical system. The target object, such as a human or boat, is very small, the interested target might occupy only several pixels or even less in one image. To detect a smalls object by applying a pattern matching technique is extremely difficult. Besides, the sun glint on the sea surface disturbs the discrimination of humans and other objects. Classification of pixels based on spectral information might be one solution, but the direct detection of the spectral signature of humans is difficult when the human wearing clothes and the spectral signature is mixed with surrounding environments. Moreover, the sunlight spectrum is not calibrated, and the background reflection varies depending on the sunlight, sea depth, etc.

In this chapter, to eliminate the variation of sunlight and background spectra, we apply partly a multispectral color enhancement method proposed by Mitsui et al.[23], in which they enhance the remaining part apart from the several significant

principal components and then merge the enhanced image to the original image. Hashimoto et al.[24] improve this method for a more effective visualization of the spectral features to the non-visible range. Their methods have successfully visualized the spectral features that need to be enhanced even when the background scene has spectral variations. We apply their method to restrain the background information. Then, for the purpose to detect small objects, we utilize a linear unmixing technique[25] to find the endmembers, and the derived endmember spectra are classified into human and other objects using spectral angular mapping (SAM) and spectral information divergence(SID) [26].

3.2 Detection Method

3.2.1 Overview of the proposed method

Figure 3.1 shows an overview of the proposed detection flowchart. We firstly normalize the data, then background information is been removed. Secondly, based on the background removed data, each endmember's spectra are been extracted and then applied different classification methods. Thirdly, a specific classification rule is defined for human detection depending on the discrimination map of each

endmember and the human spectra characteristics.

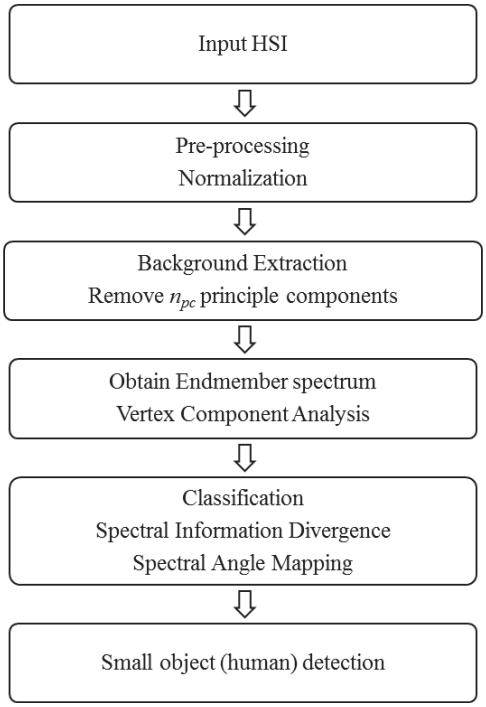


Figure 3.1 Proposed processing chain

3.2.2 Background separation

Among dimension reduction methods, Principal component analysis (PCA) has been often used. The projection data according to its contribution, the first principal component has contained the direction that data mostly distributed, the second principal component orthogonal to the first ones, and follow the direction

that data mostly distributed apart from the first principal component. The first several principal components contain most of the information on the dataset. In the case of HSI of the sea surface, they contain most of the background information.

We would like to focus on the human and other objects that suggest the presence of humans (which take only a small part of the image) by reducing the background information. We can obtain the HSI which has its background reduced by deleting the first several principal components from the original data. The HSI data cube containing H rows, W columns, and N bands (a three-dimensional data cube H*W*N), $\mathbf{s}_0(\mathrm{h},\mathrm{w})$is a column vector of the spectrum from spatial coordinates (h,w), w=1,2,…,W, h=1,2,…,H . After PCA, note $\boldsymbol{u}_\mathrm{i}$ be the *i*-th eigenvector, thus by forcing the first *n* principal components to be zero, background-suppressed spectrum $\boldsymbol{s}_{\mathrm{br}}$ can be obtained.

$$\boldsymbol{s}_{br}(h,w) = (\mathbf{0},\mathbf{0},\dots \boldsymbol{u}_{n+1} \dots \boldsymbol{u}_N)\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\\ \vdots\\ \boldsymbol{u}_{n+1}{}^t\\ \vdots\\ \boldsymbol{u}_N{}^t\end{pmatrix}\boldsymbol{s}_0(h,w) \qquad (3\text{-}1)$$

3.2.3 Endmember spectra extraction

Linear spectral unmixing methods have the assumption that a mixed pixel is

a linear combination of endmember signatures. Many algorithms have been proposed for spectral unmixing[27]. Most of the linear spectral unmixing models assume that endmembers are presented in the scene. In the HSI, at least exist one pure pixel for each endmember. Vertex component analysis (VCA)[25]is based on the convex sets. It finds the vertices of the simplex after an affine transformation. The obtained vertices are the endmembers. VCA has simpler or even better performance while at the same time its computational complexity is the lowest Compared to other linear spectral unmixing methods. Thus we apply this method to determine the endmember's spectra.

3.2.4 Classification

After the spectra of endmembers are obtained, we need to classify the HSI to detect the objects. There are many ways in hyperspectral image classifications: Support vector machine (SVM), Maximum likelihood (MLC), artificial neural network (ANN), and Spectral match classifiers such as Spectral angle mapping (SAM) and spectral information divergence (SID).

MLC has been generally applied in multispectral image classification,

however, for HSI applications the calculation time is significantly increased. ANN is similar when applied to HSI, the iteration time increased. SVM gives a promising result, although it needs to carefully choose the kernel function and other parameters. SAM and SID are very suitable for real-time proceeding system because they are easy to calculate and can be very fast.

Suppose two pixels that have spectral data:

$$\mathrm{X}=\left(x_1,\ x_2,\cdots x_n\right)^T,\ \mathrm{Y}=\left(y_1,\ y_2,\cdots y_n\right)^T$$

The spectral angle α is expressed as

$$\cos(\alpha)=\frac{\mathrm{X}\bullet\mathrm{Y}}{\|\mathrm{X}\|\|\mathrm{Y}\|} \qquad (3\text{-}2)$$

where $\alpha\in\left[0,\frac{\pi}{2}\right]$

When α is small, that shows X and Yis similar.

Spectral Information divergence, *SID*(*X, Y*) is the divergence of two spectral

to estimate their similarities, given by

$$SID(X,Y) = D(x\|y)+D(y\|x) \qquad (3\text{-}3)$$

where

$$D(x\|y) = \sum_{i=1}^{N} p_i \log \frac{p_i}{q_i} \qquad D(y\|x) = \sum_{i=1}^{N} q_i \log \frac{q_i}{p_i}$$

and

$$p_i = \frac{x_i}{\sum_{i=1}^{N} x_i} \qquad q_i = \frac{y_i}{\sum_{i=1}^{N} y_i} \quad .$$

In the proposed method, Spectra angle and SID have been calculated for each pixel compared to the endmembers obtained in section 3.2.3.

3.3 Experiment

3.3.1 Materials

A set of HSI provided by EBA-JAPAN CO., LTD has been used in this study. Images are taken using their NH series hyperspectral camera, which is a push-broom-type sensor. Each HSI is 1024 pixels (height) by 2048 pixels (width), from 400nm to 1000nm with an interval of 5nm, a total of 121 channels. Each HSI is a scene of surfboards floating on the sea with the human on or around as shown in Figure 3.1. We select a sub-region size 600 by 600 to better presented the result. The sub-region includes 11 surfboards. Some human has presented on the surfboard (the left upper corner) while some of them are not (the right downer corner), the absent human believes to exist around the board, might be in the water. Hence a total of 11 human objects need to detect in this sub-region.

3.3.2 Processing

Firstly, we extract the interested area from the original HSI data cube, then apply normalization to bring the digital value to the range [0, 1] by subtracting the

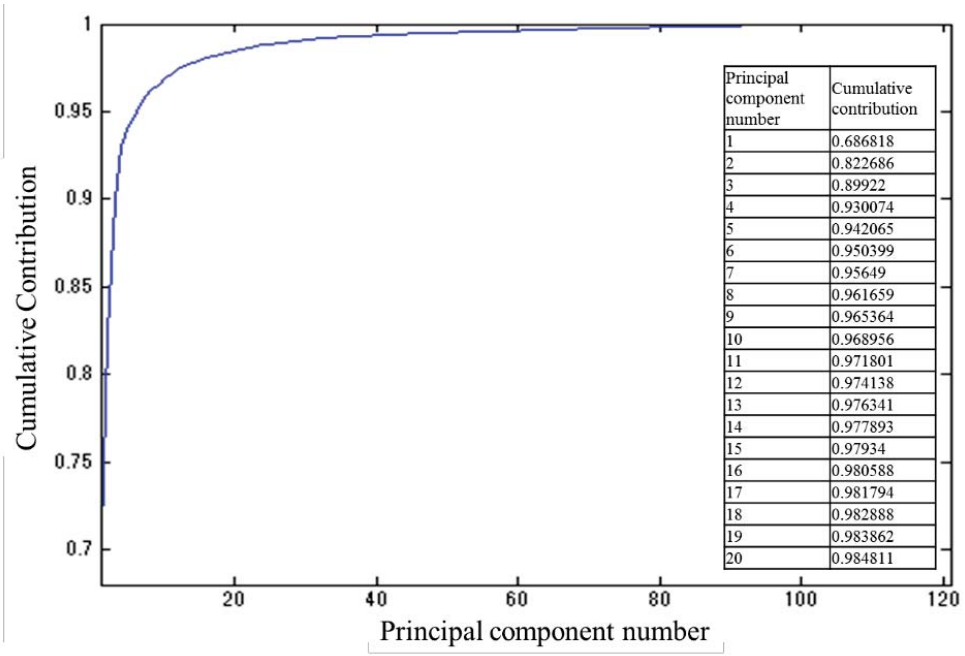
minimum digital value and then divided by the range of the digital value.

Figure 3.2 HSI used in this study, sub-region within the white rectangle is the testing area. Human positions have been marked as green rectangles.

Secondly, using equation 3-1, we remove the first n principal components. To define the value of n, the cumulative contribution of the principal component is plotted in Figure 3.3. The first 3 principal components have a cumulative contribution of about 90%, we thus define n's value for this dataset is 3.

The number of endmembers is defined as following: since there is a total of 11 surfboards in the scene, and consider the human spectrum, sea spectrum, as well

as sea glint spectrum, we define 14 endmembers to be unmixed by applying VCA.



Principal component number	Cumulative contribution
1	0.686818
2	0.822686
3	0.89922
4	0.930074
5	0.942065
6	0.950399
7	0.95649
8	0.961659
9	0.965364
10	0.968956
11	0.971801
12	0.974138
13	0.976341
14	0.977893
15	0.97934
16	0.980588
17	0.981794
18	0.982888
19	0.983862
20	0.984811

Figure 3.3 The principal component cumulative contribution

Once we obtain the endmembers' spectrum, the spectral angle and SID for each pixel compared with each endmember's spectrum to get the abundance map can then be calculated. We select 2 or 3 abundance images and set a certain threshold

to obtain the final results which present in the next section.

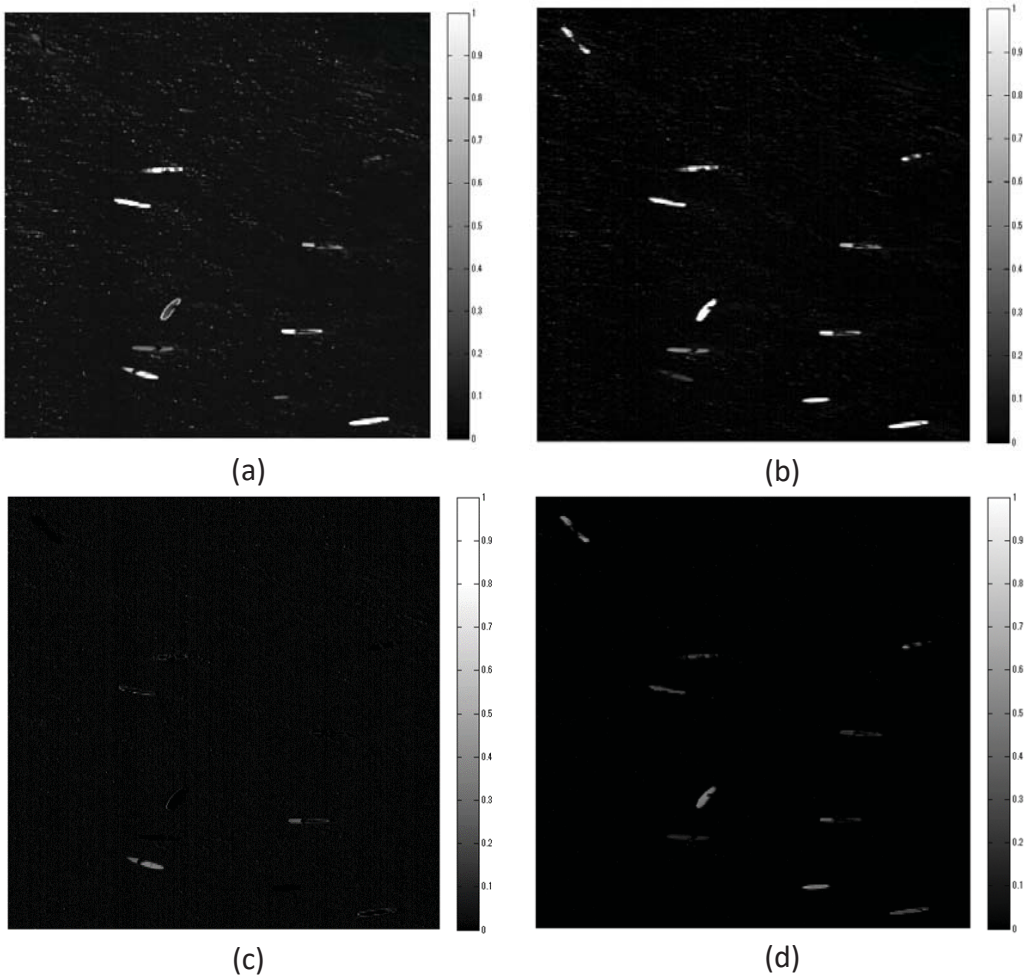


Figure 3.4 HSI images at wavelength. (a).495 nm and (b)695 nm; images after removing first 3 principal components at the wavelength (c) 495 nm and (d) 695 nm

3.4 Results

To constrain the background information, we have deleted the first 3 principal components then reconstruct our spectrum. Shown in Figure 3.4 (a) is the original HSI sub-image at 495nm, we can see the vast distribution of sun glint. Compared with (c) which is after background information constraints, we can almost see no sun glint. For 695nm is a similar situation.

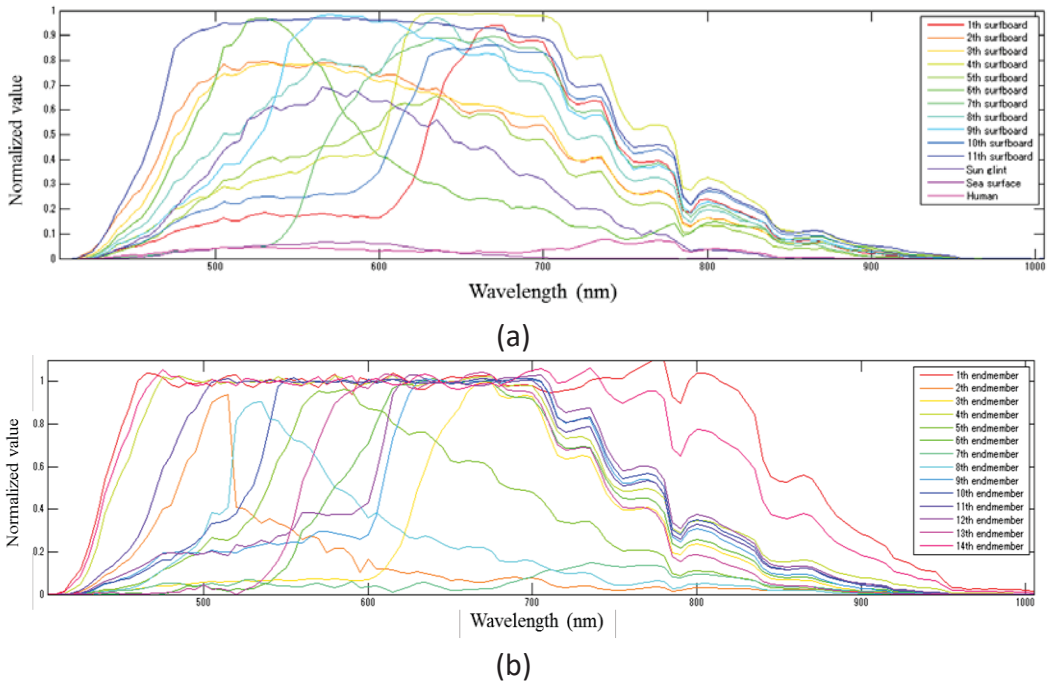


Figure 3.5 (a) Original Spectral of Objects presented in the scene. (b) Endmember obtained by VCA.

Figure 3.5 (a) shows the normalized spectral of objects presented in the

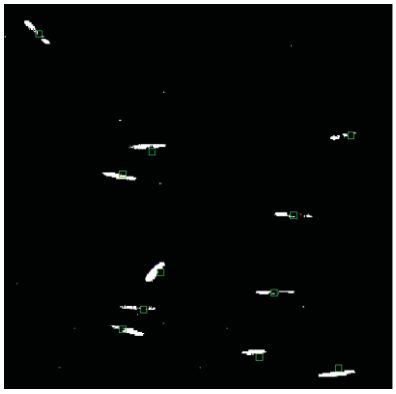
images, while Figure 3.5 (b) shows the endmember spectra results obtained by the VCA. As we can observe, the 4th endmember extracted characteristically like the Human spectra, and 9th, 10th, 11th endmember almost the same as 10th, 9th, 1st surfboard spectra, respectively. Results show the ability of VCA methods to determine endmember although some objects spectra might not be well extracted, this might be due to the saturation of data in spectra.

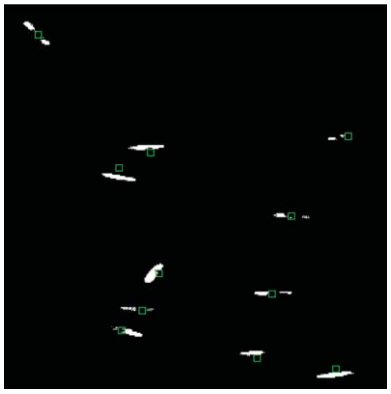
Table 3-1 Detection results (a) Without background constrain, (b) With background constrain

	Surfboard		Human		False Alarm	
	(a)	(b)	(a)	(b)	(a)	(b)
RX	11	11	4	3	12	0
SAM	9	11	7	4	12	0
SID	10	9	9	9	38	16

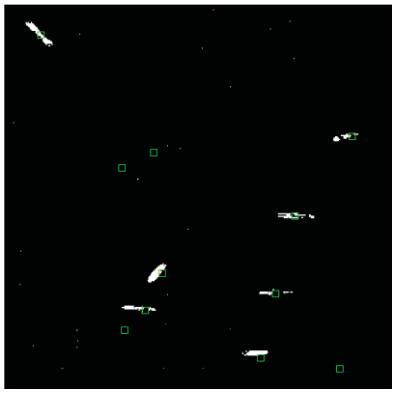
We compared our methods with and without background constrain and compare our results with the RX (Reed-Xiaoli) detecting method[24][25]. In the proposed method, SID and SAM are used in the classification step. The results have been shown in Figure 3.6.

a)

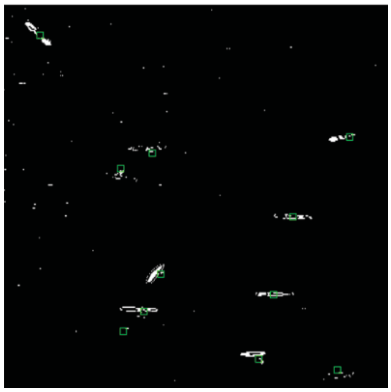




b)



c)



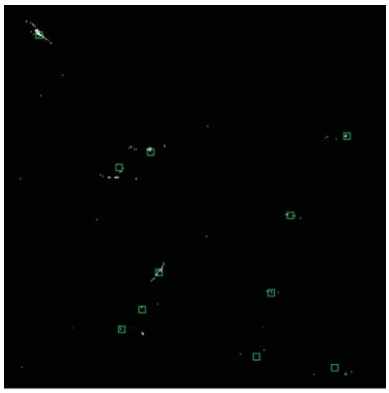


Figure 3.6. Results for (a) RX detector, (b) SAM classification, (c) SID classification, left: without removing principal components, right: remove first 3 principal components. The green rectangle is the position of human

The detection rate is shown in *Table 3-1*. Overall, the RX detector method can detect all the surfboards, however poor performance on human detection. It might cause by the human spectrum more similar to the sea surface other than the surfboard. The spectrum methods although don't detect all the boat, they perform better on human detection. This might be caused by the spectrum unmixing method that has to obtain the human spectrum signature. SID performance is better in human detection, however, at the same time, its false alarm numbers are extremely high. Therefore, it always a trade-off between a better detection rate and a higher false alarm rate.

3.5 Discussion

In this paper, we proposed a small target detection proceeding method for human detection from aerial HIS of the sea surface. That is to firstly apply background constrain and then using spectrum unmixing for classification. Experimental results suggest that the SAM method is similar to the RX detector, with an improved human detection rate. And both methods have zero false alarm after background suppression. SID method has the highest human detection while at the same time has the highest false alarm rate. However, background suppression enables to greatly reduce SID's false alarm rate, which shows the potential of using the SID method for practical human detection.

Effect of the restoration of saturated signals in hyperspectral image analysis and color reproduction

Based on Yan, L. et al.*: "Effect of the restoration of saturated signals in hyperspectral image analysis and color reproduction"*. DOI: 10.1007/s10043-020-00630-8

In hyperspectral imaging, saturation is often affected by the captured signal due to specular reflection or a peaky spectrum. In this chapter, we propose a restoration method for saturated hyperspectral signals. Our algorithm is summarized as follows: firstly, based on principal component analysis of the reconstruction basis can be obtained and a linear constrained least square problem be solved to calculate the coefficients of each basis. We discuss the problems that saturated signals might cause and apply our method to two sets of real hyperspectral images and a set of hyperspectral images with simulated saturation. Our method helps increase unsupervised object detection and improves high-fidelity color reproduction.

4.1 Background

Hyperspectral imaging (HSI) contains not only rich spectral information but

also spatial information. The spectrum signature reveals the intrinsic features of targets remotely; thus, HSI has been widely applied in different fields of research and industrial sectors, such as precision agriculture [28], geology [29], forestry [30], medical imaging [31], object detection [32], color reproduction [33], etc.

It is usually difficult for cameras to capture the full range of luminance levels of the scene as expected due to the limitation of the dynamic range of image sensors. A challenge that yields saturated signals. Signals that have reached the maximum range of the sensors and are considered a type of image distortion is called saturation, or also called clipping. This problem can also be regarded as a limitation in the amplitude of the original signal. The dilemma is how to capture both the dark and bright parts in a single shot simultaneously. Many researchers on the application of high dynamic range (HDR) imaging to hyper/multispectral image acquisition [33][34][35] have been studied; this can be accomplished by, for example, combining several low dynamics range (LDR) photos taken with different shutter speeds. Although HDR spectral imaging is possible with additional shots or sensors, it is not always employable in a number of applications due to the limitation of environmental conditions, time, and budget. Moreover, consider the saturated pixels in the hyperspectral image processing pipeline is also needed. When small, high-intensity objects are present in the scene, the exposure is usually adjusted to the

darker region to obtain a good quality image in that area. However, doing this might lead to pixels in a bright area to become saturated. Also, the possibility of signal saturation becomes more significant for higher spectral resolution in hyperspectral imaging. While for certain wavelengths already saturated. there is a higher chance at the meantime other wavelengths may not receive enough light.

Some equipment may suffer more than others due to saturation depending on the acquisition method,. There are roughly two types of hyper/multispectral image acquisition: scan and snapshot in a spectral direction. In the first approach, the device scans along a line and gradually obtains the whole image; the second method involves shooting a snapshot of the image at once. A scan type acquisition device, such as a filter-wheel camera [36], can adjust the exposure to the illumination spectrum. However, for instruments that take snapshots, such as push-broom sensors [37] or spectral filter array sensors [38], it is impossible to adjust to the illumination. The latter acquisition methods are more likely to produce saturated signals than the former one.

Severe problems in the HSI analysis can be caused by Saturated signals. Even though the spectral signatures of the target objects are not clipped, the distorted spectral data may affect the spectral unmixing, clustering, or anomaly detection. Moreover, saturation often occurs in the specular reflection on metallic objects or

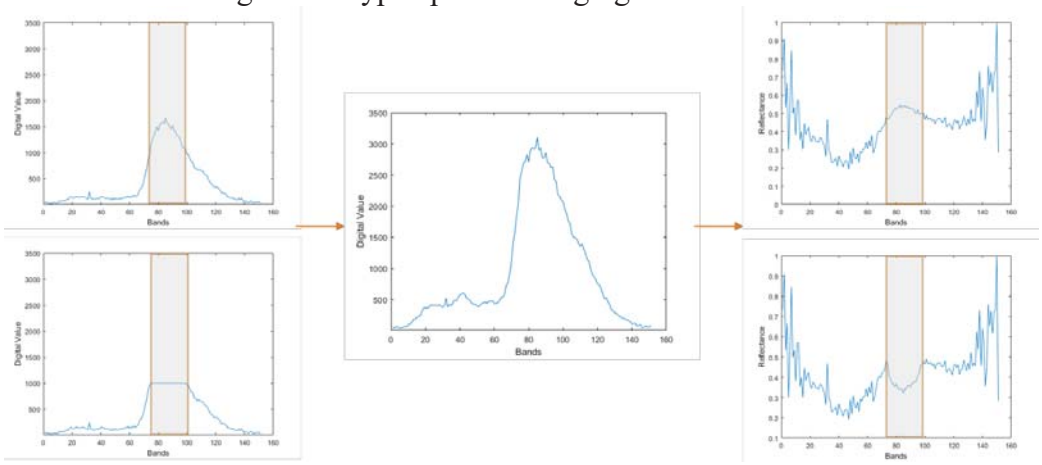
smooth dielectric object surfaces in the spectral imaging used for high-fidelity color reproduction. The color of the specular reflection or the lusters of those objects is altered if some signals in the HSI are saturated [35]. It could be considered a signal interpolation problem for restoring a clipped signal, and related work has been done in restoring digital audio signals [39][40] and computer vision [41][42]. There are some studies related to hyperspectral signal restoration, but they focus on noise reduction [43] or image deblurring [44]. To our best understanding, there is no literature on the restoration of saturated hyperspectral signals.

In this paper, how saturated signals influence hyperspectral image analysis and color reproduction have been studied. To this end, a method for restoring the saturated ranges have been proposed. The effect of this restoration is subsequently examined. The results show that restoring the saturated signal improves the detection result using anomaly detection and improves reproducing the color of specular reflection in high-fidelity color images. Section 2 presents the details of the proposed method, while the datasets and the results are shown in Section 3. Conclusions and discussion are provided in Section 4.

4.2 Methodology

In this section, we first discuss the problems that saturated signals might cause to hyperspectral imaging. We will then detail our proposed method to restore the saturated pixel.

4.2.1 Saturated signals in hyperspectral imaging



(a) Digital Values (b) Reference illumination spectrum (c) Reflectance

Figure 4.1 This figure schematically compares the reflectance calculated from a normal spectrum with the one derived from a saturated signal, which is clipped at digital value 1000. When both spectra are divided by the reference illumination spectrum, we can observe a hollow in the reflectance derived from the saturated spectrum

Saturated signals refer to the pixels or areas that have been wrongly recorded, an effect resulting from the limitation of the sensor's sensitivity range. Improper exposure time or other kinds of mistakes will also lead to saturation, this kind of

manual mistake will not be considered in our study. We will only focus on the inevitable saturation case.

In natural scenes, the light is usually not balanced throughout the whole image: some parts of the scene might be darker while other areas are much brighter. It is a tradeoff between exposure time and preserving details in a low light area. A short exposure time will avoid saturation; however, darker parts of the scene might not have enough light. At the same time, a long exposure time will better represent the more shaded areas while suffering the risk of saturation. Light-emitting objects, fluorescence, or the specular reflection of the light source can as well cause saturation.

When saturation happens, the signal is partially lost, thus causing problems in hyperspectral imaging applications. For example, when calculating the reflectance of a spectrum, if there are saturated signals within certain bands, then the reflectance of those bands will be calculated to be smaller than the actual reflectance, as illustrated in Figure 4.1. Those parts will tend to produce false colors accordingly when applied to color reproduction under different illuminants. If this happens in anomaly detection, spectra that are only a constant multiple of other spectra should usually be considered normal. In that case, a saturated spectrum would have a different shape and thus be wrongly counted as an anomaly. Furthermore, if we

apply normalization to the spectral data, the spectrum will have different distortions depending on the magnification.

4.2.2 Restoration of saturated hyperspectral signals

In the proposed method for restoring saturated hyperspectral signals, we employ principal component analysis (PCA). PCA uses a linear transformation to project the original data into another feature space, in which every basis is orthogonal to each other; these bases are called principal components (PCs). By their contributions, Principal components are ranked, and the top ones contain most of the information of the projected space. PCA and PCA-based methods have been widely used in hyperspectral image processing for dimension reduction, feature selection, classification, object detection, color technology [45], etc.

There are assumptions that the spectral reflectance of real-world objects can be represented by a linear combination of several basis functions [45]. Since the spectrum at position (x, y), $f(x, y, \lambda)$, captured by HSI, is the product of illumination spectrum and the spectral reflectance. In our case, the saturated pixels only consist of a limited part; their contribution can be neglected if we use a small

number of basis functions, the result is not affected even if they are not excluded.

As shown in Figure 4.2 for each pixel that contains any saturated wavelength, our algorithm applies three steps, as follows.

Step 1: We find the saturated part in this spectrum, which means the value reaches the maximum value of the sensor.

Step 2: We calculate the coefficient of each PC only for the unsaturated bands; for a pixel that contains a saturated wavelength range, the real spectrum $\widehat{f_t}(x, y, \lambda)$ can also be described using the same basis functions, as

$$\widehat{f_t}(x, y, \lambda) = \sum_{i=1}^{n} \beta_i(x, y) U_i(\lambda) + \hat{\varepsilon}(x, y, \lambda), \qquad (4\text{-}1)$$

where $U_i(\lambda)$ is the i[th] principal component basis function, $\beta_i(x, y)$ is the coefficient for the pixel at position (x, y) and $\hat{\varepsilon}(x, y, \lambda)$ is the error. Then, we estimate $\beta_i(x, y)$ by the following equation:

$$\hat{\beta} = \underset{\beta}{argmin} \left\| \widehat{f_t}(x, y, \lambda_{ns}) - f(x, y, \lambda_{ns}) \right\|_2^2, \qquad (4\text{-}2)$$

where

$$\widehat{f_t}(x, y, \lambda_{ns}) = \sum_{i=1}^{P} \hat{\beta}_i(x, y) U_i(\lambda_{ns}), \qquad (4\text{-}3)$$

and λ_{ns} is the non-saturated wavelength, $\hat{\beta} = \{ \hat{\beta}_i(x,y) \mid i = 1 \dots P \}$ is the set of estimated coefficients, P is the number of basis functions used in the estimation, $\widehat{f_t}(x,y,\lambda_{ns})$is the estimated spectral within non-saturated wavelengths, $f(x,y,\lambda_{ns})$ is the original spectrum within non-saturated wavelengths and $\|.\|_2$ is the L-2 norm.

Since the saturation signal should always above the saturation value (S), we added the following constraint:

$$\sum_{i=1}^{P} \hat{\beta}_i(x,y) U_i(\lambda_s) > S, \ \lambda \in \lambda_s , \qquad (4\text{-}4)$$

where and S is the saturation value, and λ_s denotes the saturated wavelength. Taken together, Equations 4-2, 4-3, and 4-4 constitute a constrained linear least square problem. We use the interior-point method (a default function of the constrained linear least square solver in MATLAB) to find $\widehat{\beta}$ for λ_{ns}, where the coefficients are calculated from the non-saturated band. band.

Step 3: Finally, we consider that the value of $\hat{\beta}$ derived from the non-

saturated wavelength range is equal to the saturated wavelength, which yields

$$\hat{f}(x,y,\lambda) = \begin{cases} f(x,y,\lambda)\,, & \lambda \in \lambda_{ns} \\ \sum_{i=1}^{P} \hat{\beta}_i(x,y) U_i(\lambda), & \lambda \in \lambda_s \end{cases}, \qquad (4\text{-}5)$$

Note that the original signal remains the same for the non-saturated wavelength λ_{ns}, while the signal at the saturated wavelength λ_s is reconstructed by the estimated spectrum. This may cause a discontinuity at the boundary between the saturated and unsaturated bands; but since the constrained least-squares method has been used to calculate the coefficients, the reconstructed spectrum in the unsaturated bands becomes close to the original spectrum. This kind of discontinuity is then negligible. Still, one should notice that in certain circumstances, such as if a large percentage of the image has low signals (zeros everywhere), our proposed method might perform poorly.

The number of basis functions used in the restoration is determined by a

cumulated contribution larger than 0.985, which is empirically defined.

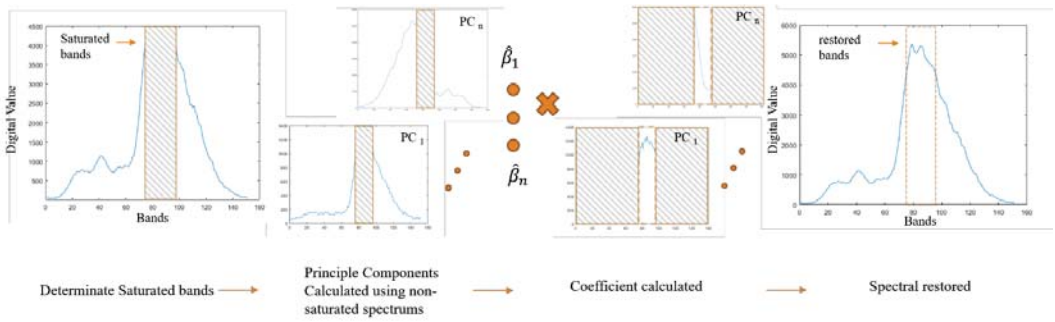


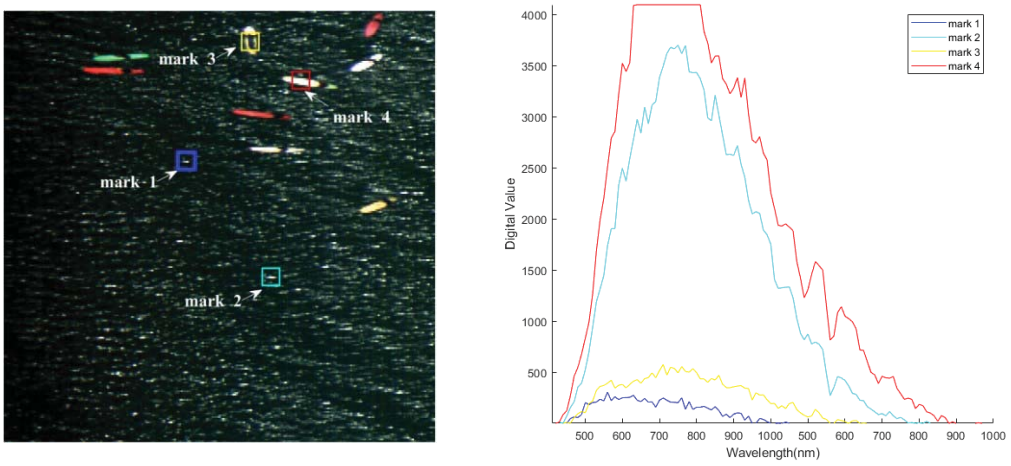
Figure 4.2 The flowchart of our method.

4.3 Experimental Results

Two experiments by applying our proposed method have been carried out. The first example is to increase the detection results in the anomaly detection problem. Another one is applying color reproduction; the results indicate that our method can improve color reproduction in specular parts. Our method is also applied to a simulated hyperspectral dataset by setting a virtual saturated digital value to a real dataset, results show our method can restore the signal properly.

4.3.1 Object detection

We used a dataset collected using the NH series camera provided by EBA JAPAN [46]. This dataset contains 43 hyperspectral images with 121 bands, ranging from 400 to 1000 nm with an interval of 5 nm. The data was captured above a sea surface; the scene contains the sea surface, several surfboards, and people. The images are stored on 12 bits; that is to say, the camera's gray level is from 0 to 4095 for each wavelength. We apply the dataset to find the people on the sea surface. Note that there are many sun glints on this surface, as shown in Figure 4.3 Mark 1 and mark 2 represent the sea surface, with limited sun glints and numerous sun glints, respectively; mark 3 represents a person on a surfboard, while mark 4 the surfboard's reflection. Reflections of sunlight are usually saturated.



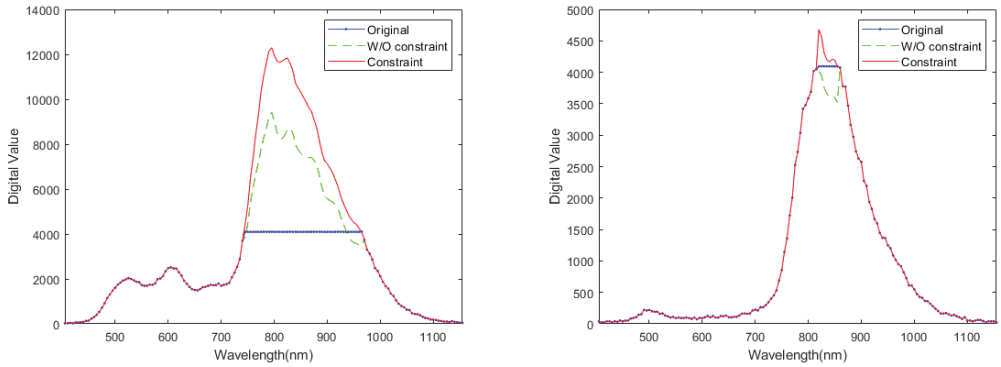
(a) (b)

Figure 4.3 (a) Examples of airborne HIS of the sea surface. Surfboards and humans are observed on the surface. (b) the average spectral data picked from marks 1 – 4 in the image (a).

We have applied the proposed method to this dataset. Figure 4.4 shows the plot of an example of the original saturated spectrum (blue dash-dot line), the restored spectrum without the constraint (the green dash line), and with the constraint (red line). Without the constraint, the reconstructed signal might sometimes drop below the saturation value, which is not logical since if it is not saturated, it should be recorded correctly. Moreover, the spectrum does not change in the non-saturated wavelength range.

The Reed-Xiaoli anomaly detector (RXAD) [4] is a benchmark anomaly detection method; under the assumption that the background is homogeneous and can be represented with a multivariate Gaussian distribution. A global RXAD has been chosen among the different types of RXADs, known as the Mahalanobis distance detector. Each test pixel is calculated using the Mahalanobis distance to the mean of the background vector. Those pixels that have a distance greater than a pre-defined threshold are considered as anomaly pixels. This detector has been applied to the original and restored hyperspectral images, and the results are presented in Figure 4.5. The red dots mark the regions detected as anomalies. Anomaly pixels are those that do not belong to the background (sea surface), such as surfboats, humans, etc. The RXAD method yielded a lot of anomalies, which include all the objects we want to detect and a substantial amount of false alarms at the same time. This might

be caused by the fact that the background is complicated, and the distribution is influenced by the anomalies. Therefore, it is crucial to reduce false detection.

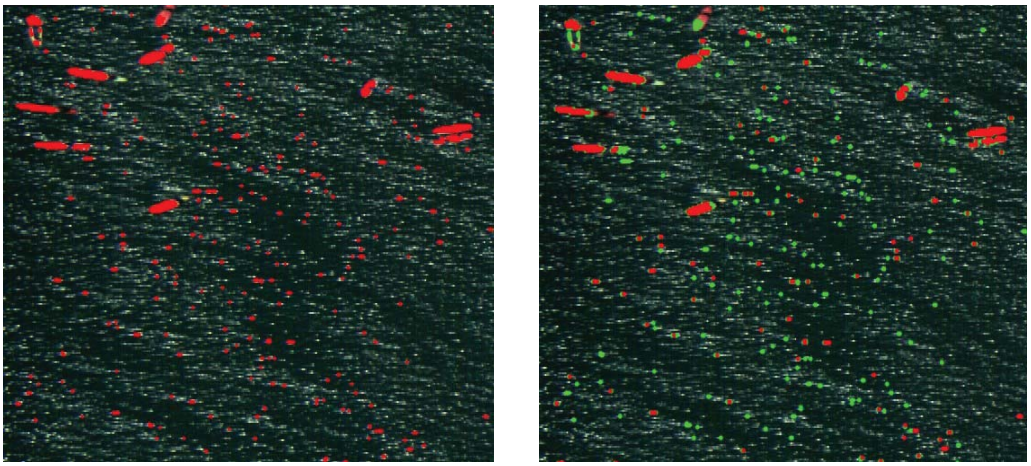


(a) (b)

Figure 4.4 Selected examples of spectra. The blue dash-dot line shows the original spectrum, the solid red line the one with the constraint, and the green dash line the one without the constraint.

On the sea surface area of the original HSI, the RXAD detected many sun glints as anomalies. The number has decreased significantly while after the saturated spectra were restored. We manually masked all the objects (surfboards and humans) in the image as ground truth for anomalies. The objects were grouped by 8-connected pixels, while each anomaly pixel was grouped by an 8-connection to its surrounding anomaly pixels. If an anomaly hits the object area, we count it as a true positive (TP), and one that misses all the objects is considered a false positive (FP); we compared the TP versus FP detection results over 43 images. Our method cannot

increase the number of TPs since the RXAD has detected all of them (therefore, there is no room to improve). After restoration, the FPs have been reduced, while the number of TPs remains almost the same or slightly decreases, as shown in Figure 4.6. The reason for this is that saturated spectra in the sun-glint areas have been restored and are therefore not considered anomalies.



(a) (b)

Figure 4.5 Object detection results using the Reed-Xiaoli anomaly detection (RXAD) method. (a) RXAD on the original data and (b) on the restored spectrum. The detected pixels are denoted with red, while the differences in detection are shown in green in (b).

In our previous work, we also applied spectral unmixing to distinguish humans from false detections[19]. Among spectral unmixing techniques, we studied vertex component analysis (VCA) [25] since it has the advantage of lower computational complexity compared to other unmixing methods. VCA assumes the data can be considered a linear combination of pure spectra, with the constraint that

the abundances are always positive. It determines a predefined number of pure components (also called endmembers) by finding the vertices of a simplex of the data.

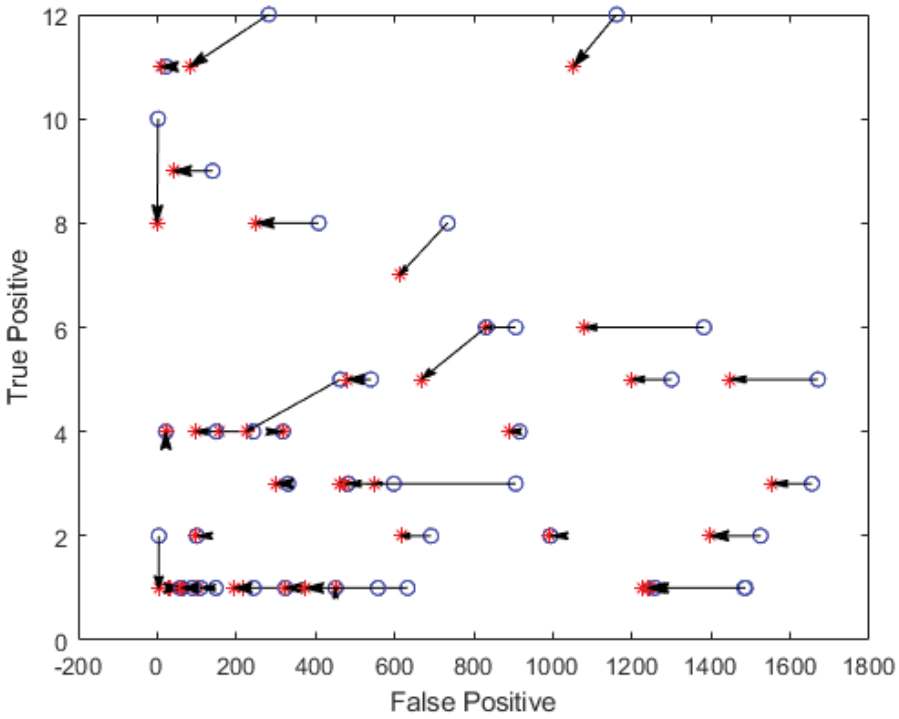
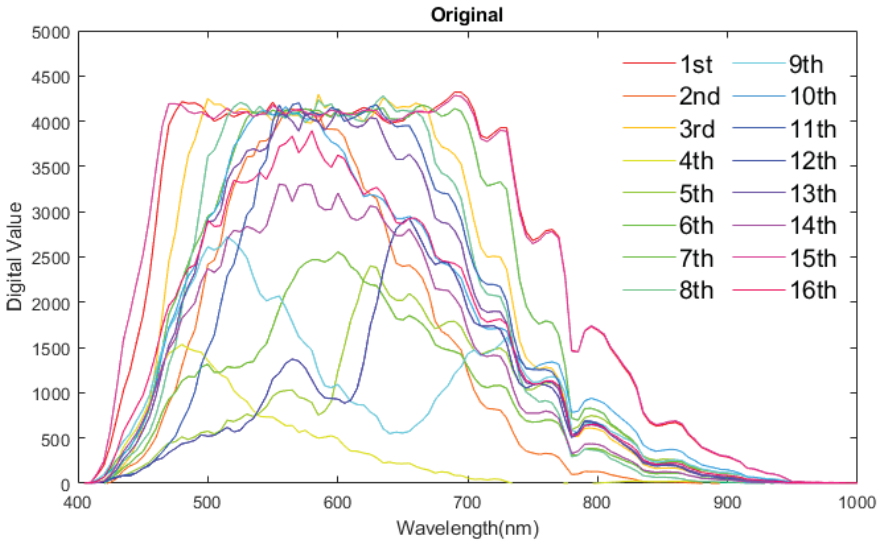


Figure 4.6 True positive and false positive comparison for 43 images. The blue circles represent the detection results for the original data, while the red stars the results using the restored data.

Figure 4.7 shows the endmembers extracted from the original and restored datasets using VCA. In Figure 4.7 (a) we can find that several endmembers have strange flat signatures. Instead, this phenomenon was solved in the restored dataset,

which is more physically consistent with the object spectral signatures.

(a)



(b)

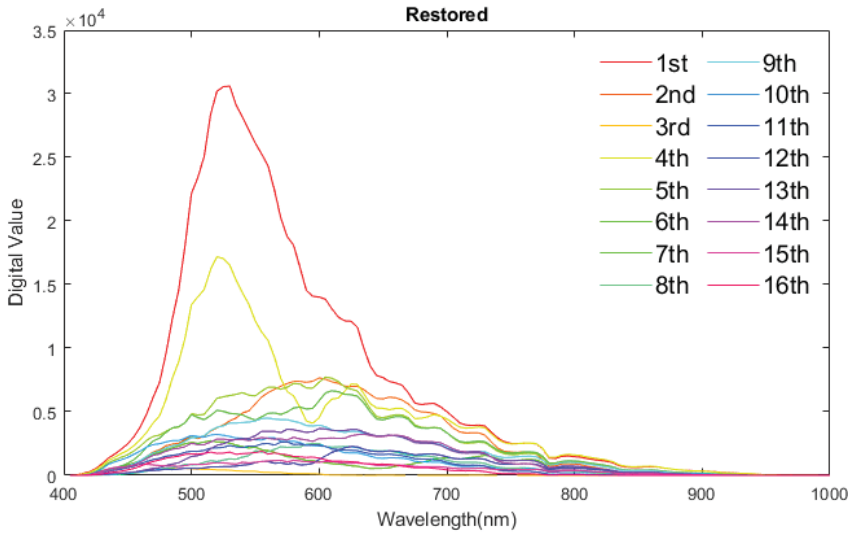


Figure 4.7 Endmember extracted (a) from the original image and (b) from the restored image

4.3.2 Color reproduction application

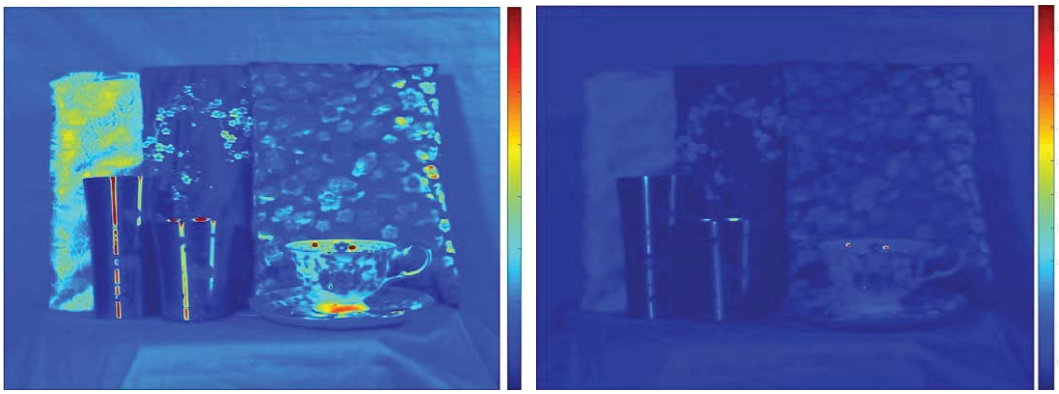
Spectral imaging has been used for color analysis and reproduction. Using a spectral image, it is possible to reproduce the high-fidelity color of a scene as well as to simulate the color appearance under different illuminants [47]. Moreover, many applications—art paintings [48], food quality [49], pathology [50], etc.—use spectral imaging for color analysis.

(a) (b)

Figure 4.8 (a) Saturated versus (b) restored RGB images. Note that the color range in the (a) saturated image was adjusted such that the color of the non-saturated region appears the same in both images.

One of the difficulties in color reproduction is specular reflection, and its reproducibility affects the impact on the realistic impression of the scene and the material appearance of objects. Specular reflection usually has a luminance about ten times higher than that of a white diffuse surface under the same illuminant and thus is likely to be saturated. Once the pixels of specular reflection saturated, the

color of the gloss is lost and the impression of the image is affected. Concerning dielectric objects, specular reflection has the color of the illuminant according to the dichromatic reflection model [51]. In the case of metals, the color of the specular reflection is determined by the Fresnel reflection formula and depends on the type of metal.



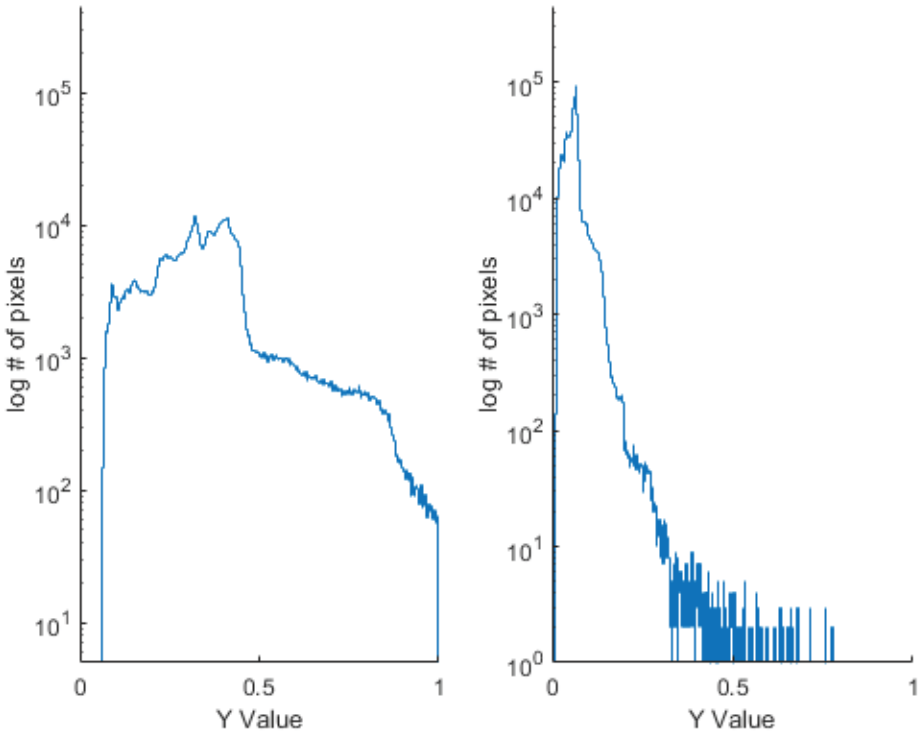
(a) (b)

Figure 4.9 (a) Saturated vs. (b) Restored Y channel images(before normalization). Note that the range of color assignment is different.

We set up the scene for color reproduction: two metallic thermal bottles (a tall and a short one) without lids and a teacup with a saucer were placed in the front. In the back, there were three Japanese brocades with gold threads(Figure 4.8) The brocades have a unique luster due to the fine texture of the gold thread, while high-luminance lusters were caused by the direct reflection of the source lights on the surface of the two thermal bottles and the brim of the teacup. We have used two halogen flood lamps to illuminate the scene with color filters of 5900 K. The

hyperspectral dataset contained 151 bands, ranging from 350 nm to 1100 nm, with a step of 5 nm. Only the visible range was used in the color reproduction. The dataset was obtained with an NH series camera (EBA JAPAN Co. Ltd.) [19]. The 12-bit gray-level image on each channel had a resolution of 1024×1280 pixels. The scene was captured along with a white reference under the same exposure time. The reflectance of the scene was calculated by dividing every pixel value by the spectrum of the white reference. The CIE standard color-matching functions are used to calculate the XYZ values under the D65 light source, values which were converted to an RGB image based on the sRGB standard. The Y values were normalized to those of the white reference, which made the Y values of diffuse objects smaller than one. At the same time, specular reflection gave rise to saturated pixels, for which the Y values were larger than one. Figure 4.9 shows the luminance (Y) channel of the original and restored images. The original image range is about 1.5, however, it increased up to 6 after the restoration. In the original image, many parts (the ones marked in red) reached maximum luminance; however, in the restored image, the luminance of a large fraction of such areas was reduced to the average level, with only two points on the rim of the teacup remaining red. Please note that to calculate the RGB value, the luminance has to be normalized first. In Figure 4.10 the luminance is normalized. In Figure 4.10 which is the histogram of the two images on the Y channel, the distribution has been changed. The increasing pixels with a

higher luminance in the restored image means the dynamic range of the Y channel was improved.



(a) (b)

Figure 4.10 Histogram of Y channel of the (a) original saturated and (b) restored images. The horizontal axis shows the luminance value, and the vertical axis indicates the number of pixels in the log scale.

Four parts from Figure 4.8 were enlarged and shown in Figure 4.11 (a)–(d) to closely compare the results for the original saturated images (first columns) and restored images (second column). The third and fourth columns plot the RGB pixel values along the cross section indicated by the red arrows. The horizontal lines

(panels [a] and [b]) are plotted from left to right, while the vertical ones (panels [c] and [d]) from top to bottom.

More details can be observed after restoration. Figure 4.11 (a) shows how the fine texture on the gold petal, indicated by the yellow arrowhead, appears flattened in the saturated image. In Figure 4.11 (b), the teacup's gold rim is white, while after being restored, its actual gold color (indicated by the lower arrowhead) becomes visible. The gold round pattern between the two four-leaf flowers on the inside of the teacup (indicated by the top arrowhead). The metallic luster in Figure 4.11 (c) appeared as pure white in the original image, on the contrary, the restored image shows more details on the decoration lines of the thermal bottle. In Figure 4.11 (d), the silk pattern looks slightly brighter and perceptually better.

Considering that the color of specular reflection represents a property of the material, the proposed method is valuable for reproducing a more faithful representation of textures. In this paper, we illustrate a simple but effective way to improve the impression of lusters.

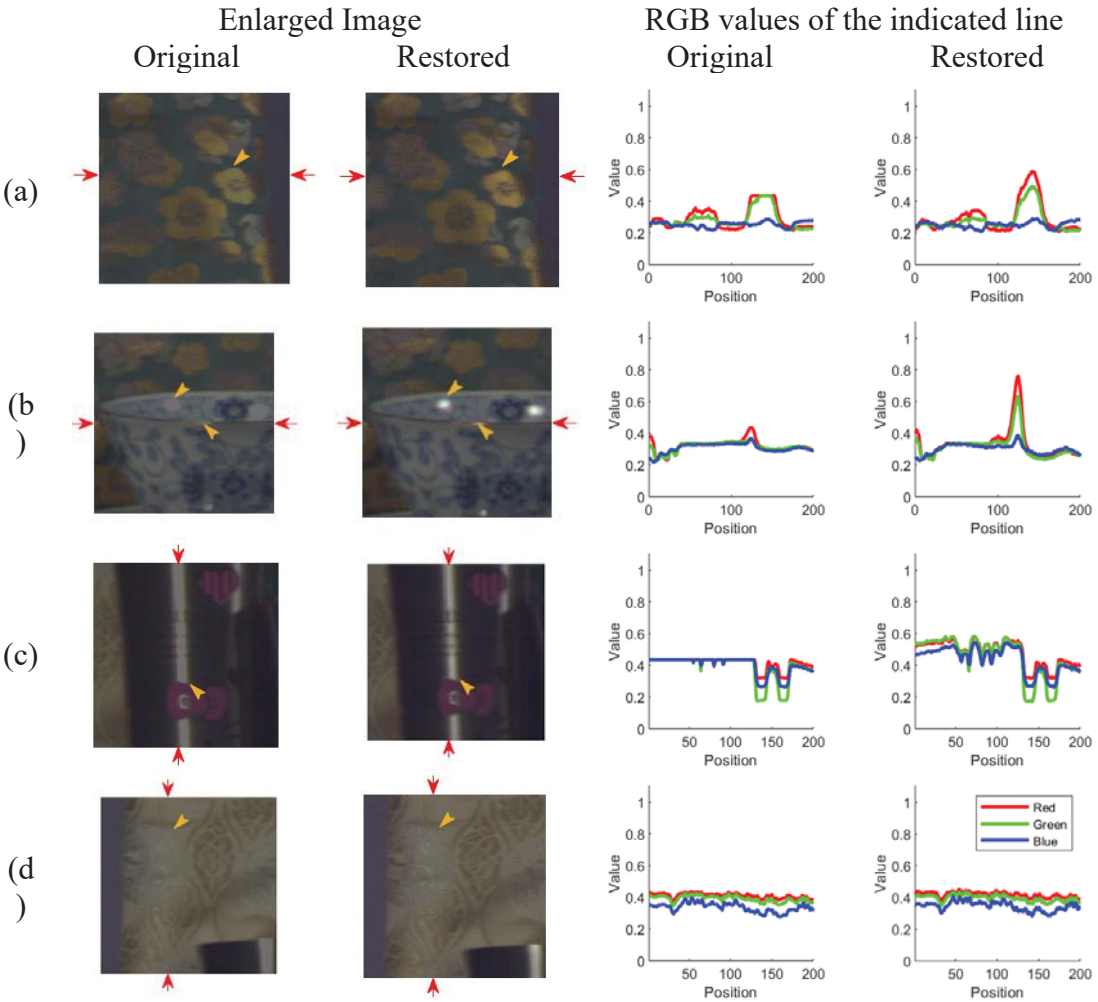


Figure 4.11 (a)–(d) Enlarged parts of the original and restored RGB images and the profiles of the RGB values before and after the restoration along the lines indicated by the red arrows.

4.3.3 Simulated saturated hyperspectral data in color reproduction application

To quantitatively and colorimetrically evaluate our method's performance, we performed the following experiment, in which we simulated the saturation of hyperspectral data.

Four patches were selected in which exist almost no saturation pixels but still some highlight pixels. Then, by forcing the pixels with the spectrum larger than the threshold reduced to the threshold, we simulated the signals taken by a lower dynamic range device. By applying this thresholding, the spectrum may acquire some saturated bands; the original data was the ground truth. We then compared our restoration results with the ground truth and evaluated the performance of the proposed method.

While the original data's maximum value was 4095 since the images are stored on 12 bits, the maximum value was set to 2047, assuming an 11-bit storage device. In the simulated captured images in Figure 4.12, the impression of the highlighted pixels is different from that in the original images. In the third column of Figure 4.12, our method successfully improved the appearance of the glossy or highlight regions.

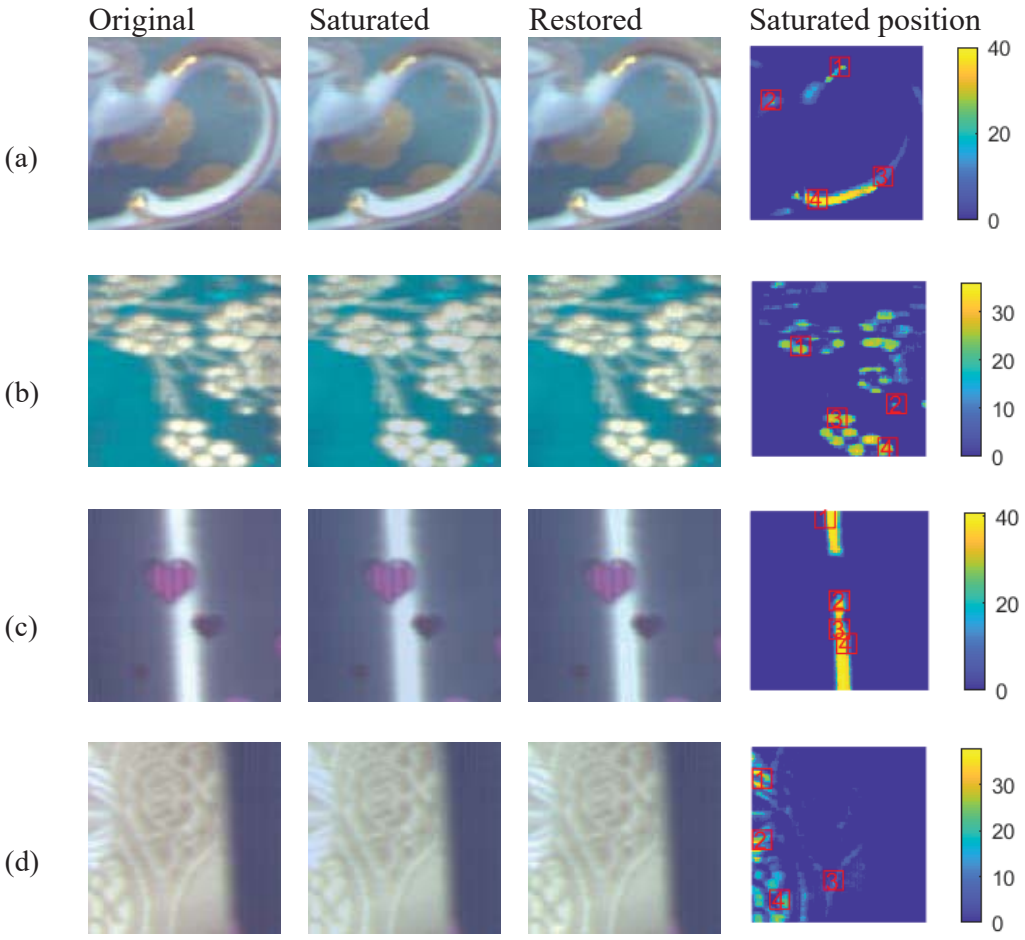
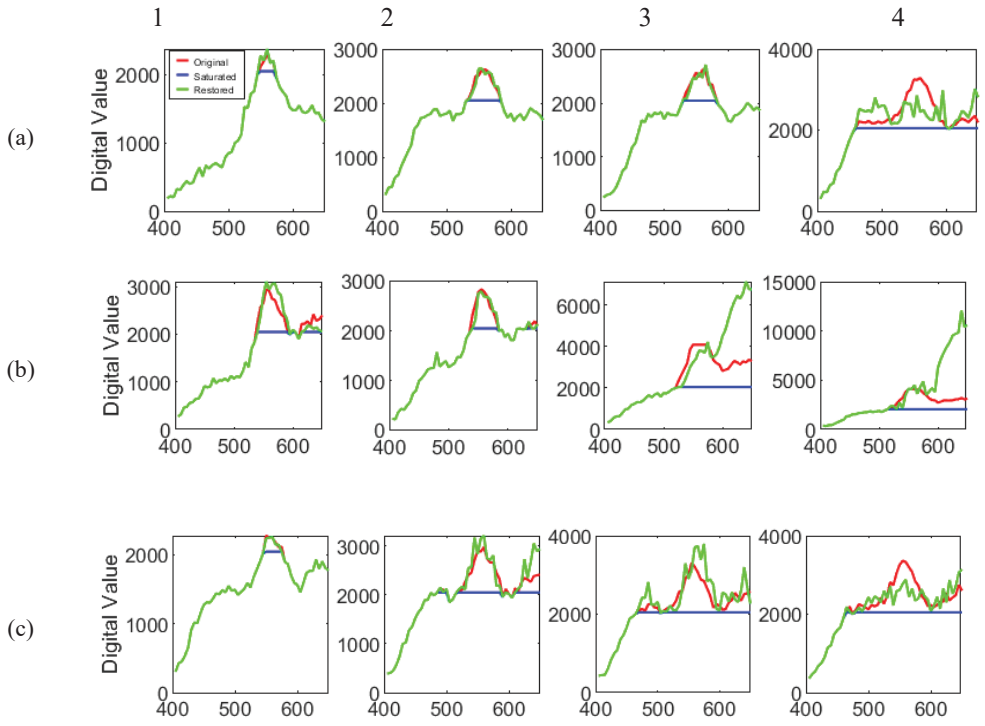


Figure 4.12 RGB images of original data, saturated data, restored data, and the saturated positions (the false color represents the total number of bands saturated in the spectra) where the spectral data were sampled.

The sampled spectral data of the original, saturated, and restored images are

shown in

Figure 4.13. Each spectrum is obtained from the single pixel centered in the square shown in the saturated position column in Figure 4.12. It can be confirmed that the restored spectra resemble the original ones in most cases. Though, the failure cases have shown in some instances (see a-4, b-4, and c-4) where the spectrum reached the maximum value for a broad continuous range of wavelengths.



(d)

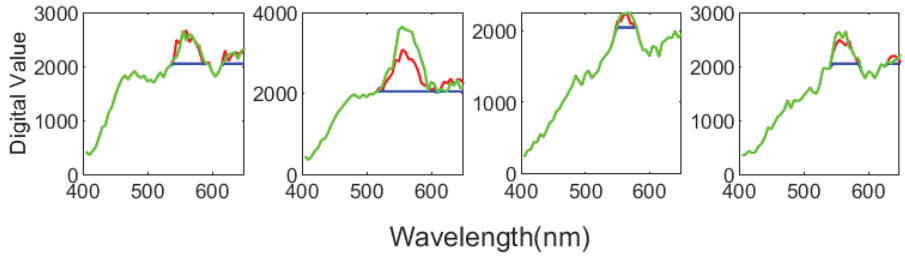


Figure 4.13. The corresponding spectra of different sub-areas: the original (red), saturated (blue), and restored (green) signals. The labels a-1 ~ 4, b-1 ~ 4, c-1 ~ 4, and d-1 ~ 4 indicate the pixels' positions in Figure 4.12, wherefrom the spectra are obtained.

To colorimetrically and quantitatively evaluate the results, we compared the CIELAB ΔE. As a matter of fact, the CIELAB color space is not ideal for evaluating the color differences between highlighted pixels, but ΔE can give a certain extent valuable information on how the color is reproduced. The results are shown in Figure 4.14, where one can observe that ΔE is smaller in many pixels after restoration. Nevertheless, we also found that our method failed for some pixels, which had a

large ΔE even after restoration.

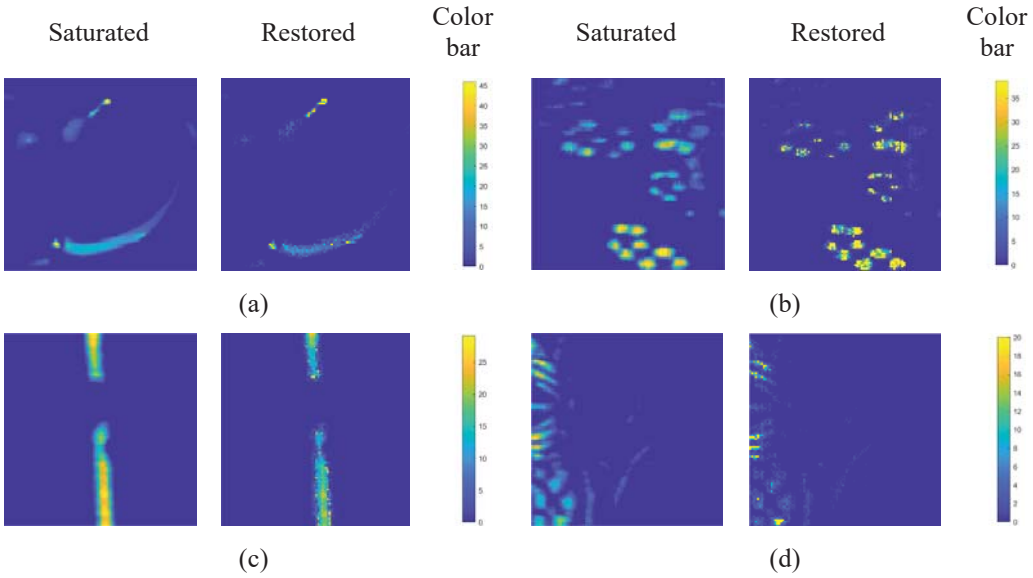


Figure 4.14. CIELAB ΔE of saturated (left) and restored (right) patches.

We observed the color gamut in the $L^* - C^*$ plane where chroma $C^* = \left(a^{*2} + b^{*2}\right)^{1/2}$ for each patch, as shown in Figure 4.15 to further analyze the effect of the proposed method. The results demonstrate that our method increased the lightness range compared to the saturated images. The distribution became similar to the ground truth. The maximum L* values in the saturated images were about 90, whereas the restored image had pixels with L* ~ 100. Still, we noticed some points with much higher chroma than the original after applying our method. Figure 4.16 shows the detailed spectra of those points with the largest chroma value (C* value). The restored data had much higher values than the original in a longer wavelength

range. These errors were also caused by the broad range of continuous saturated bands. The information in the non-saturated bands was not satisfactory for estimating the saturated wavelength range. Nevertheless, the color difference is not evident, although perceivable. When observing the restored images in Figure 4.16, It seems that high chroma values in L* >100 are not too colorful, and the results of the proposed method still appear acceptable.

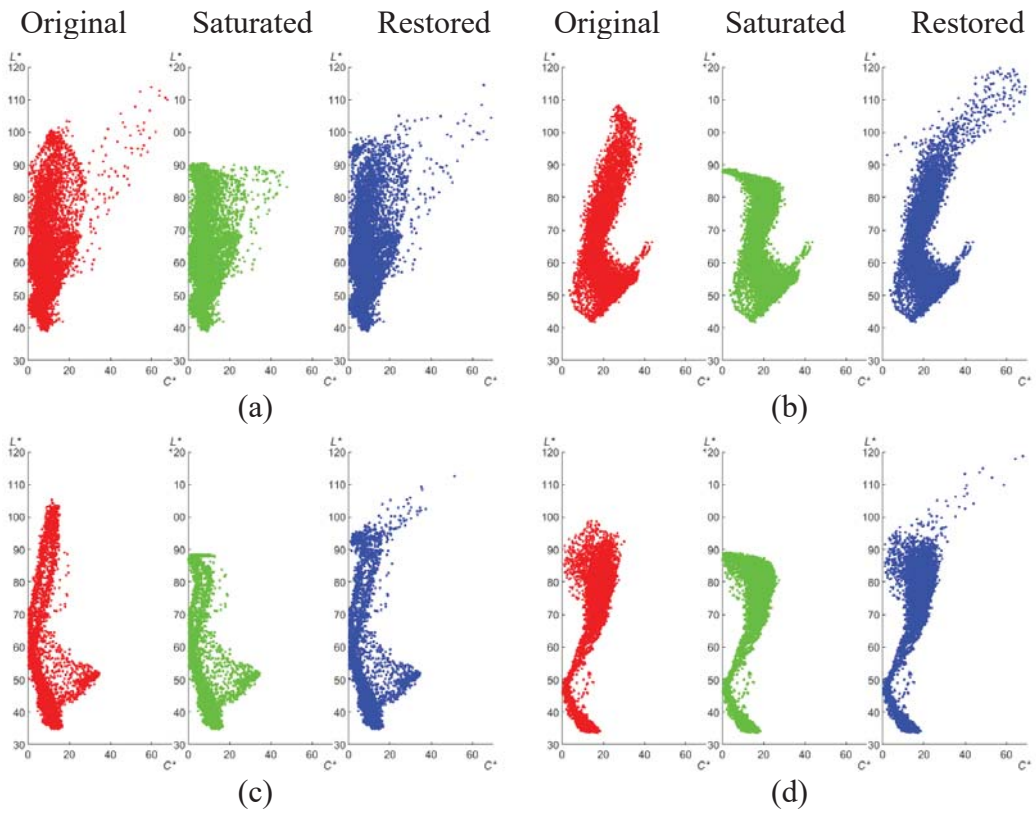


Figure 4.15 Color gamut: the distribution in $L^* - C^*$ plane for each patch corresponding to the patches (a)-(d) in Figure 4.12

These results suggest that our method fails to work if the ratio of saturated bands is high—where the constrained least square equation gives an invalid solution—and thus we cannot restore the saturated bands. For this reason, we must specify that the proposed method is effective under the circumstance that the ratio of saturated bands in a pixel is low, a limitation of the proposed method. A solution to this problem could be using constraints on spatial continuity or spatial distribution. Furthermore, it would be valuable in the future to apply machine learning to the restoration of highlighted features by training the spectral and spatial

characteristics of specular reflection or emission in various scenes.

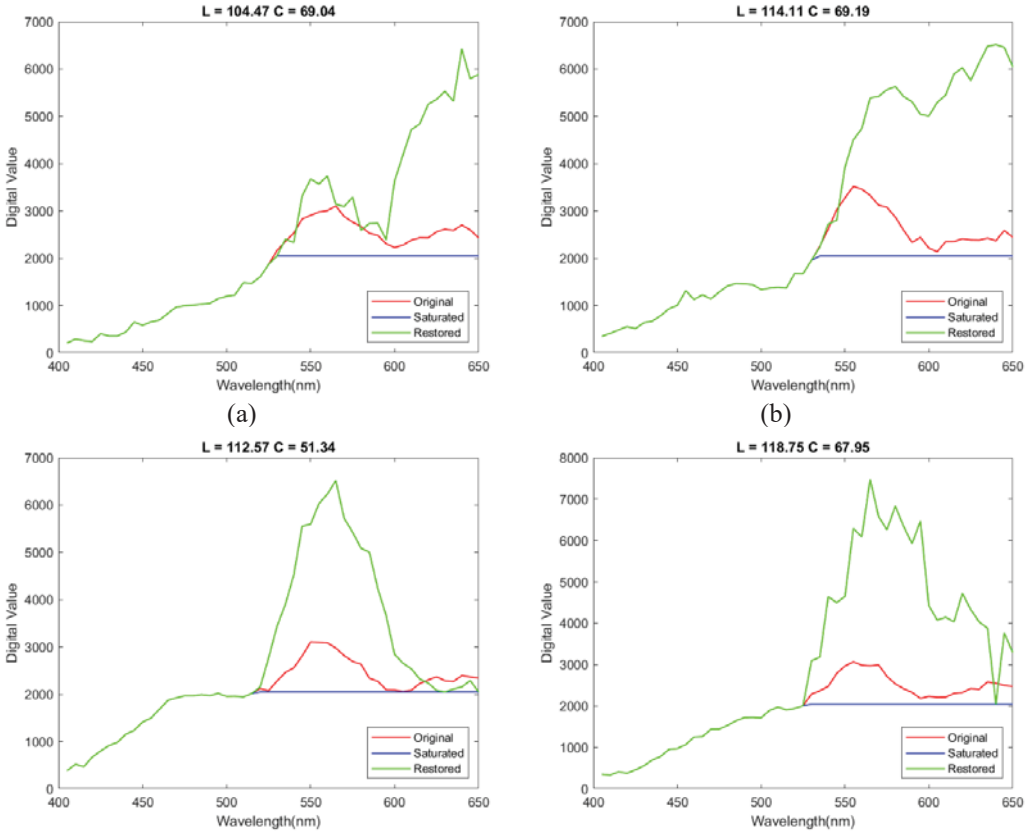


Figure 4.16 The original (red), saturated (blue), and restored (green) spectra of each rightmost (highest chroma) point shown each patch in Figure 4.15

4.4 Conclusion

A hyperspectral saturated signal restoration method has been proposed., under the assumption that the spectral reflectance function of an object is a linear combination of several basis spectra presented in the same scene. For illustration, we applied only PCA to demonstrate the idea that a restored image could bring

benefits to object detection and color reproduction. This method could improve the accuracy of unsupervised anomaly detection after obtaining the results of real HSI. Yet, our approach is limited by the linear combination assumption. A large part of the image is saturated will cause our method's underperformance. In future work, we could also investigate more advanced algorithms (instead of PCA) for restoration, such as using sparse optimization, deep learning architecture, or spatial/spectral statistics. Our method also shows the potential to help increase color reproduction in the glossy regions of an image.

Method 2: two-stage deep learning-based method

Based on Yan, L. et al.: "*A novel two-stage deep learning-based small-object detection using hyperspectral images*." DOI: 10.1007/s10043-019-00528-0

Hyperspectral imaging's application to object detection and classification currently becomes an important research topic. However, accurately identify small-sized objects remains challenging. In this second, we proposed a novel two-stage deep learning-based hyperspectral neural network(2SHyperNet) suitable for human detection from the sea surface. Spatial and spectral information of hyperspectral images has been combined. We use pixel-wise spectral information in the first stage to obtain 1st-stage classification results and then combined the result with spatial information to help eliminate unlikely regions. In this way, the detection accuracy has been improved. We tested our method on a dataset of real-world airborne hyperspectral images. And we compare the performance with several conventional methods. The proposed method outperforms current state-of-the-art methods.

5.1 Background

Spectral imagery has been widely used recently for the sake of its advantage: it contains rich spectral and spatial information. Compare to traditional RGB images, spectral images contained much more information by acquiring data at a different wavelength. And hyperspectral imaging (HSI) is containing hundreds of narrow band images. Hyperspectral images have the advantage of revealing the nature of the object because the spectral information of an object has its unique characteristics. HSI has been firstly used in remote sensing for Earth studies such as mining and geology [52]. With the development of HSI, its application has been extended to various domains such as precision agriculture [53], archaeology and art conservation[54], forensic science[55], medical diagnosis [56], and so on.

Detection and classification of desired objects are one of the most important applications of HSI. Statistical signal processing and machine learning (ML) algorithms suitable for HSI have been developed[57] such as support vector machine (SVM)[58][59], random forest (RF)[60] and deep learning (DL) [4][61][62][63][64][65][66], etc. SVM has been widely used in classification problems since its efficient, good classification result and does not easily cause overfitting. But it has some drawbacks: 1) the learning speed is usually slow due to the kernel function evaluations, especially in the presence of noise and a large amount of training samples; 2) the parameters needed to be tuned, which depend

strongly on the user's experience. RF classifier is an ensemble classifier that consisted of multiple decision trees. A subset of training samples and variables are randomly selected. The advantage of RF classifier is a high accuracy, fewer parameter, and good performance on high dimensional data. Nevertheless, it might be suffered by overfitting and sensitivity with noisy data. DL has shown its ability for image recognition, object detection, and classification with RGB images, and has been further applied in hyperspectral data processing and classification problems. Makantasis[66] have used a typical conventional neural network (CNN) after dimension reduction of HSI. But, CNN based-method applied in object detection usually does not perform well in small-sized object detection. This is caused mainly by two factors, one is in convolutional layers, feature maps have been down-sampled many times, the other is in the pooling layer, spatial information can be lost if the stride is greater than one [67]. X. Ma et al, have proposed a spatial updated deep auto-encoder network using for spectral-spatial classification which shows deep learning could extract spectral and spatial information at the same time[68].

In most of the research which focuses on the classification of hyperspectral images, their training data amount needs to be balanced on the same scale. Manolakis et al.[69] have mentioned that the classification problem aims to minimize the probability of misclassification errors while the object detection

problem is to seek to maximize the probability of detection while keeping the probability of false alarm under a certain predefined value. A small-sized object detection problem is to detect those objects only take a very little part of the image. One of the difficulties in small object detection is that objects take part in only a very limited part of the image. For example, to detect some small-sized objects from the background, suppose such objects take only 1% in total of the image, even a method has class such object wrongly as the background class, such method will still get 99% overall accuracy. The problem that needs to solve is how to detect the small-sized object from the non-homogeny background.

An anomaly detector can distinguish targets whose signature is spectrally different from the surrounding background. Reed and Yu have proposed an anomaly detector called Reed-Xiaoli Detector (RXD)[4]. This detector measured the Mahalanobis distance of the processing point's spectral and the average spectral of its surrounding window. Wei Li et al. trained a CNN which learn the similarity between spectrum and then use such a network to detect anomaly in the HIS[70]. In section 2 we will introduce more related methods.

Back to the beginning of 21 centuries, most of the hyperspectral images in remote sensing are acquired by satellite, luckily, with the development of hyperspectral cameras, aerial images' requirement has been increasing. The aerial

image has the advantage of personalization and flexibility compared with the satellite image. Small object detection could be improved because of the reduction of atmospheric distortion and optical distortions have been reported in recent works[71].

An application example of the hyperspectral small object detection is the human detection from airborne HSI as shown in Figure 2.2 [19]. As the UAV flying over the searching area, an alarm will send to the user once the potential target has been found. The user can check or verify if the detection is valid by zooming in or paying more attention to the alarm areas after receiving the alarm.

5.2 Hyperspectral Small-sized object detection

Small-sized object detection methods can be categorized as supervised, unsupervised, and semi-supervised depending on the requirement of training data. The method, basic assumption, advantage, and disadvantage are summarized as follows.

Supervised methods are those which need training data, or we at least know

prior knowledge about the spectral characteristics. By learning the features or distribution, we will be able to separate the target from the background. An unsupervised method for small-sized object detection also known as anomaly detection (AD), which does not need any training data. For hyperspectral imagery, RXD is known as the benchmark AD [16]. Semi-supervised methods are those which need limited priority knowledge of the target or background as training data. However, they do not need as much training data as the supervised method.

To conclude, the supervised method needs a relatively large amount of training data as well as good quality of training data, and this kind of method also has a high risk of overfitting. But they will have higher accuracy if the testing data have the same distribution of training data. The unsupervised and semi-supervised method needs user interpretation of the results and might have low detection accuracy for a specific target.

Thus, as we aim at high accuracy, in this paper we propose a supervised deep learning-based method. But to be noticed, our proposed methods only need to provide limited training data of target, the negative samples will be selected from the background. Our proposed method can prevent overfitting by combining spatial

and spectral information.,

5.3 Proposed Deep Learning architecture

Hyperspectral image data has the advantage of containing both spectral and spatial information. In deep learning classification of hyperspectral data, spectral information is contained in each pixel along the different wavelengths, and meanwhile, the spatial information is in the image taken at each wavelength.

To better use both spatial and spectral information, a two-stage deep learning architecture has been proposed, as shown in Figure 5.1: the first stage performs learning a rough spectral feature of the dataset, while its primary classification results are combined with the spatial information (using convolution layer to learn the spatial distribution) in the second stage to get more precious detection.

In our study, each hyperspectral data cube is C(columns)×R(rows)×B(bands). For the stage 1 neural network (S1NN), the sample is a vector $\boldsymbol{x} = (x_1, x_2, \cdots, x_b)^T$, and we have a total $N_{total} = C \times R$ samples for one image. The output of the first stage is spectral classification probability $\widehat{y_i} =$

$f_{stage1}(\boldsymbol{x}_i), i = 1,2,\cdots, N_{total}$. In this stage, we have two dense layers ($D^i_{stage1}, i = 1,2$), with 64 and 32 units respectively, and we use a rectified linear unit (ReLU) as an activation function for these two dense layers. ReLU layer applies the function $Relu(x) = \max(0, x)$ in all elements on an input tensor, this layer can reduce the computational complexity, captures non-linearities, and reduce gradient vanishing problem. Suppose in each dense layer, $\boldsymbol{W}^i_{stage1}, i = 1,2$ is the weights and $b^i_{stage1}, i = 1,2$ is the bias. Then the S1NN can be expressed as:

$$f_{stage1}(x|\boldsymbol{W}, \boldsymbol{b}) = \hat{y} \qquad (5\text{-}1)$$

Then we use the Sigmoid (Logistic) function, shown in (2), to predict the probability of the target object. We choose binary cross entropy (BCE), shown in (3), as the cost function since it's a two-class classification problem. Adaptive Moment Estimation (Adam) has been choosing as an optimizer because Adam adds bias-correction and momentum and might be the best overall choice according to[67].

$$Sig(z) = \frac{1}{1+e^z} \qquad (5\text{-}2)$$

$$BCE = -\frac{1}{N}\sum_{i=1}^{N}[y_i\, log(\hat{y}_i) + (1 - y_i)\, log(1 - \hat{y}_i)] \qquad (5\text{-}3)$$

The convolutional layer (CONV) and dense layers are utilized in the Stage 2

Neural Network (S2NN). CONV produces feature maps accordingly to each filter that have been defined by users with specific size and numbers. In our architecture, we use 2D CONV (Spatial CONV) with size (N×N) and a 1D CONV (Spectral CONV) with size (1×N). The S2NN is composed of two parts, where the first part extracts the spatial feature of the original data, this part is composed of 5 sets of layers: the first set is composed of one Spatial CONV with 8 filters of size (4×4). After that follows one batch normalization (BN) layer, the BN layer normalizes layer inputs, this layer will force your activations to be zero mean and unit standard deviation and will reducing gradient diffusion thus can prevent overfitting. The next layers are one ReLU layer with one dropout layer, the following 4 sets of convolution layers are composed of one Spectral CONV layer, one BN layer, one ReLU layer one dropout layer with the first two sets have 16 filters (4×1) and last two sets have 32 filters for CONV layer.

In the second part of S2NN, which starts with a flatten layer to change the shape of the data into a vector, then the vector can be inputted to a fully connected layer. The neurons are concatenated with the flattened results from the S1NN. The concatenated vector is then fed to a dense layer followed by a dropout layer and two dense layers with 128 units with ReLU as the activation function. The last dense layers have 1 unit and the sigmoid function is used in the last dense layers for

probability prediction. BCE is used as the cost function.

Our proposed 2-stage architecture has the following advantages. Firstly, the S1NN only needs a small amount of training data. This can reduce the labeling efforts, especially for hyperspectral image labeling which usually takes more effort than original RGB images. Secondly, as the S1NN take only spectral information, it might have the risk of overfitting, in S2NN, it will take the S1NN results as a clue for the potential candidate, however at the same time, S2NN will learn the spatial distribution of the target thus to help reduce the overfitting and also can reduce the false detection. Thirdly, once the neural network has been trained, the prediction process is not time-consuming, thus might have the potential for task-specific real-time detection.

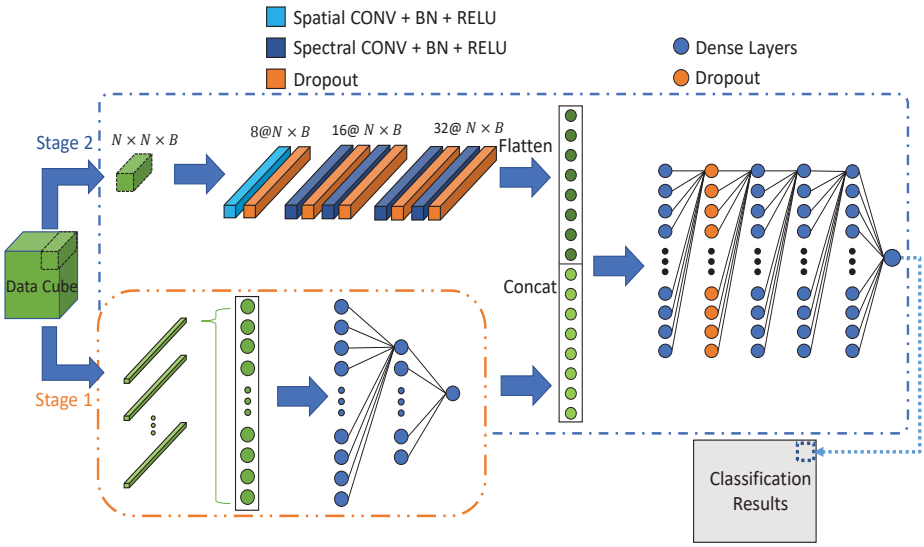


Figure 5.1 Proposed deep neural network structure

5.3.1 Dataset

A set of HSI data are taken using the NH series hyperspectral camera[46]. Each HSI is in 512 pixels (height) by 1024 pixels (width), spectral range from 400nm to 1000nm with an interval of 5nm, a total of 121 channels. Thus, the captured hyperspectral images are of the size 512×1024×121, Each HSI is a scene of surfboards floating on the sea with humans on or around as shown in Figure 5.2

As we can observe, although these images have been taken on the same day and the same area, there is great variation among those images we can observe in Figure 5.2. The depth of the sea, the angle of the sunlight, and the angle of the camera

towards the scene will all influence the appearance. We select 36 images out a total of 60 images that contain at least one human object. The locations of humans and surfboards are manually marked as ground truth as in Figure 5.2 (d). The human label is considered as positive samples and the others are considered are negative samples. We use all the human label samples in training images and select the same amount randomly from the surfboard samples and sea surface samples as the negative samples since the human label sample is limited. The number of negative samples is about twice as positive ones. Figure 5.3 shows the spectral variation in different regions: sun glint, human, surfboard, and sea surface without sun glint. In each image, hundreds of spectral have been shown together, the solid line shows the mean of those spectral while the color background shows the maximum and minimum of the spectral. As demonstrated in Figure 5.3 we can find that the spectral signature seems very different between classes and meanwhile there is also variation within the same class.

5.4 Experimental Results and Comparisons

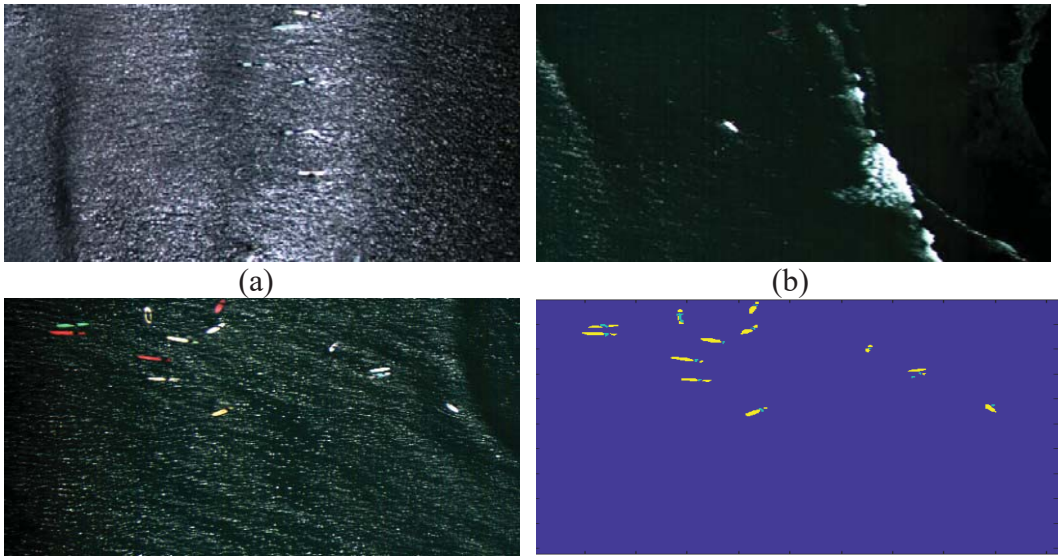


Figure 5.2 (a)(b)(c) is a sample image from our dataset, (d) is the ground truth of (c) with yellow represents the surfboard and the light blue represents the human, the dark blue is the background sea surface

To compare the performance of the proposed method, we perform the experimental verification, where the methods for comparisons are selected as the linear SVM, RF, Makantasis[66]'s CNN, stage 1, and stage 2 only respectively. The experiments are implemented with Keras[72] using TensorFlow[73] as a backend. All experiments were performed on Intel i7-4790K CPU and NVIDIA GTX TITAN X GPU.

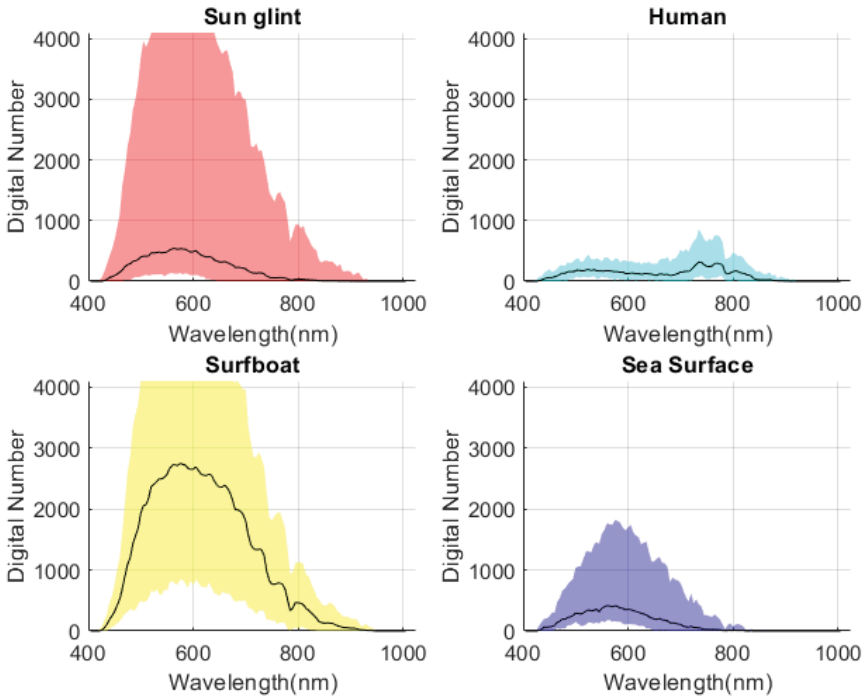


Figure 5.3 Spectral Signature sun glint, human, surfboard and sea surface without sunglint

We used leave-one-out cross validation method to test different methods' performance, thus the data was divided into 36 folds. In each fold, training data was from 35 images and 1 image was left out for evaluation. For the linear SVM, RF we used Scikit-learn toolbox[74]. Every sample was a 121-dimensional vector. For the linear SVM, we used Tolerance of 0.0001 and maximum 10000 literation, output was the probability, and the other parameters were set as default values. For RF, we built a model with 300 trees in the forest, with a max depth was 4, the other

parameters were default. Makantasis's CNN was the reimplemented version[75]. The input was a hyperspectral data cube with the position of the target, the other parameters were default. To train our NN, the input for S1NN was 121-dimension vectors. For the S2NN was a 4 × 4 × 121 data cube. The batch size of the training data was 64, and it trained for 20 epochs.

For quantitative assessment, we used Receiver Operating Characteristic (ROC) curves, which indicate true positive rate (TPR) to the false positive rate (FPR), with TRP and FPR defined as:

$$TPR = \frac{TP}{TP+FN} \quad (5\text{-}4)$$

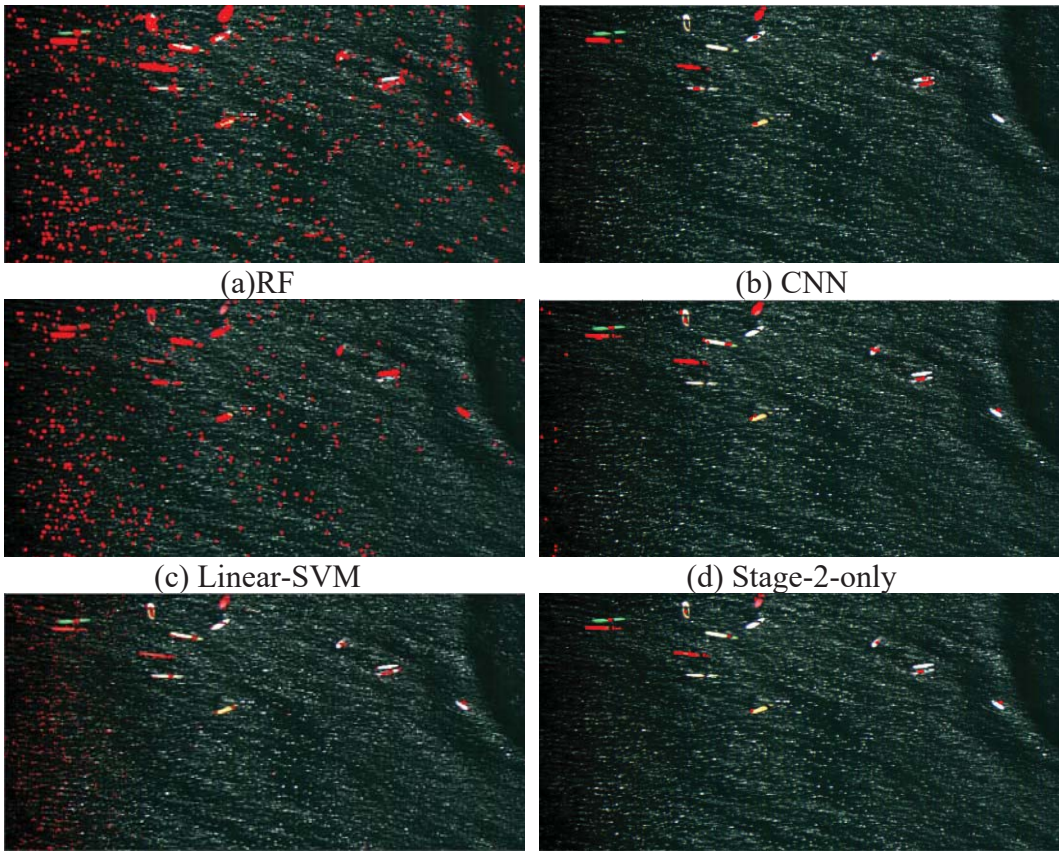
$$FPR = \frac{FP}{FP+TN} \quad (5\text{-}5)$$

where True Positive (TP) represents the number of correctly detected human pixels given a certain threshold; the sum of TP and False Negative (FN) represents the total number of the human pixels in the image. False Positive (FP) represents the number of pixels detected as humans from the background; and the sum of FP and True Negative (TN) is the number of background pixels. The accuracy of each method depends on how well an algorithm separates the human target from the background, and it can be measured by the area under the curve (AUC) of the ROC

curve. The higher the area is, the more accurate.

Figure 5.4shows the target discrimination results produced by our proposed method as well as compared methods. As we can see, in the results of RF, linear SVM, and stage 1 only, many false detections on the sea surface were classed as human. They also detected the surfboard as human. This is because those methods only used spectral information. In contrast, the methods that used spatial information (the CNN, Stage-2-only, and proposed method) show less misclassification on the sea surface. On the other hand, the CNN performed very poorly in human detection; this is because the spatial information have been lost on the convolution layers since, at these layers, the filters tended to learn the background features. Stage-2-only method performed better than the CNN, but it also suffered from some misclassification. Our proposed method's detection results were more accurate than other conventional methods: it well detected the human position while remaining

low false alarm rate.



(a)RF (b) CNN

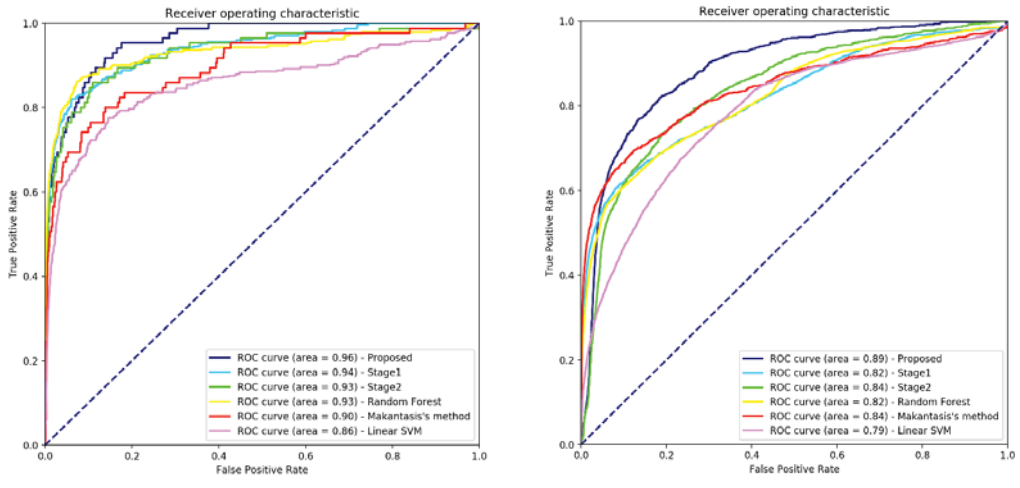
(c) Linear-SVM (d) Stage-2-only

(e) Stage-1-only (f) Proposed

Figure 5.4 Detection result on one image

Figure 5.5(a) illustrates ROC curves for the demonstrated image, using the proposed algorithm as well as compared methods. Our proposed method has AUC of 0.96 while stage 1 and stage 2 only method has AUC of 0.94 and 0.93. Moreover, as we observed when the TPR reach to 1, our proposed method showed the lowest

FPR. We also compared the leave-one-out cross validation on 36 images, and the ROC curve is shown in Figure 5.5 (b). The overall AUC is 0.89 for the proposed method, which is better than all the other methods. When reaching to TPR rate at 0.9, the FPR of the proposed method was below 0.16, and the proposed method performed significantly better than other methods. In human rescue applications, it is very important to avoid overlooking the target human. The proposed method enables to reduce of FN cases without increasing FP, which is an important advantage of the proposed method. On the one hand, S1NN helps us to focus on the highest possible candidate, which results in reducing the FN, while on the other hand, S2NN rule out those candidates have different spatial distribution.



(a) (b)

Figure 5.5 ROC curve (a) for the demonstrate image (b) leave one out cross validation on 36 images

For the human rescuing purpose, it's very important to not miss the potential target while remaining a low false detection. As we take a close look at the details of the detection as shown in Figure 5.6, Makantasis' method and linear SVM almost missed all the humans presented, while the RF detected the human however with many false detections. S1NN's results were slightly better but also had a lot of false detections. Our proposed method well detected the position of the human while with much fewer false detection shows that the performance of the proposed method is better in human detection on the sea surface with a relatively low false alarm.

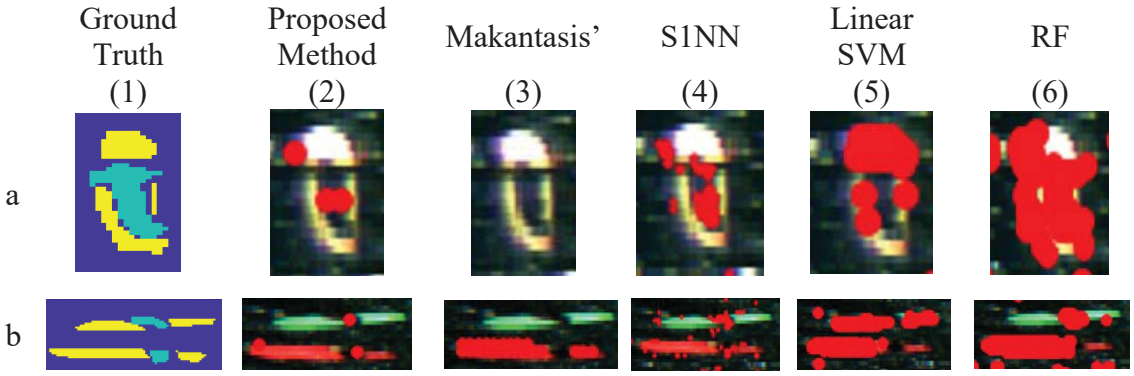


Figure 5.6 Details on human detection performance. Ground truth is shown in column (1), as yellow represent the surfboard, dark blue represents the sea and light blue represent the human.

5.5 Conclusions

Object classification using HSI has been widely studied but little work on small-sized object classification has been reported. In this paper, we proposed a 2-

stage deep learning method for human detection from aerial HSI of the sea surface. Our approach constructs high-level features that encode pixels spectral and spatial information. Our method has been compared to the Linear SVM method and RF method as well as only use separately stage. Experimental results show that our proposed method presents superior performance compared to the others. This method might inspire this approach to be applied in other HSI small object detection problems.

Conclusion

In this thesis, we investigate state-of-the-art small-sized object detection methods and propose two methods of using the hyperspectral dataset to perform human detection from the sea surface. The first method is semi-supervised and requires human interaction when the system provides a potential target spectrum. This method does not need training data and can thus also be useful in other applications of small-sized object detection. However, this method suffers from a relatively high false alarm rate since it only uses spectral information. To better detect a human object from the sea surface, we then applied a deep learning-based method. We proposed a two-stage neural network to combine spatial and spectral information. Results show that this structure can give a particularly good detection rate while retaining a low false alarm rate. While processing the hyperspectral signal, we found that by restoring the saturated bands, the anomaly detector can work better and can create better fidelity in color reproduction. This thesis might inspire the future approach to be applied in other HSI small object detection problems, such as bruise detection on fruits, disease detection on leaves, small-sized abnormal

detection in various medical images, etc.

In Chapter 3, we proposed a small target detection proceeding method for human detection from the aerial HIS of the sea surface. That is to first apply background constraint and then use spectrum unmixing for classification. Experimental results suggest that the SAM method is similar to the RX detector, with an improved human detection rate. Both methods have zero false alarms after background suppression. The SID method has the highest human detection rate and the highest false alarm rate. However, background suppression enables it to greatly reduce SID's false alarm rate, which shows the potential of using the SID method for practical human detection.

In Chapter 4, we proposed a hyperspectral saturated signal restoration method. This approach was designed under the assumption that the spectral reflectance function of an object is a linear combination of several basis spectra presented in the same scene. For illustration, we applied only PCA to demonstrate the idea that a restored image could bring benefits to object detection and color reproduction. The results of real HSI show that this method could improve the accuracy of unsupervised anomaly detection. However, our approach is limited by the linear combination assumption, and if a large part of the image is saturated, our method might not be able to perform well. In future work, we could also investigate

more advanced algorithms (instead of PCA) for restoration, such as using sparse optimization, deep learning architecture, or spatial/spectral statistics. Our method also shows the potential to help increase color reproduction in the glossy regions of an image.

In Chapter 5, we proposed a two-stage deep learning method for human detection from aerial HSI of the sea surface. Our approach constructs high-level features that encode pixel spectral and spatial information. Our method has been compared to the RF and linear SVM methods, as well as using only a single stage. Experimental results show that our proposed method presents superior performance compared with the others.

In our study, the training data is limited and collected in a short range of time, so the light can be considered as varying on a small scale. However, in a more common situation, the light in the scene can vary. Future work can develop a high-performance detection method that can tolerate light variation. An example is a model trained with sunny day hyperspectral data but also performs well on cloudy days. One possibility is to apply few-shot learning or zero-shot learning, which has been recently widely applied in object detection and semantic segmentation using

RGB images.

With the limit of hyperspectral spatial resolution, pixels mixed with multiple spectral are quite common in the scene. This will lead to difficulty and inaccurate segmentation. In future work, super-resolution methods might overcome such sub-pixel segmentation problems and thereby increase detection results.

References

[1] Du, B., Zhang, Y., Zhang, L., & Tao, D.: Beyond the sparsity-based target detector: A hybrid sparsity and statistics-based detector for hyperspectral images. IEEE Transactions on Image Processing, 25(11), 5345-5357. (2016)

[2] Wang, T., Zhang, H., Lin, H., & Jia, X. A Sparse Representation Method for a Priori Target Signature Optimization in Hyperspectral Target Detection. IEEE Access, 6, 3408-3424. (2018)

[3] Park, J. J., Oh, S., Park, K. A., Foucher, P. Y., Jang, J. C., Lee, M., ... & Kang, W. S.: The Ship Detection Using Airborne and In-situ Measurements Based on Hyperspectral Remote Sensing. Journal of the Korean earth science society, 38(7), 535-545(2017)

[4] Reed, I. S., & Yu, X.: Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. IEEE Transactions on Acoustics, Speech, and Signal Processing, 38(10), 1760-1770(1990)

[5] Salem M.B., Ettabaa K.S., Hamdi M.A.: Anomaly detection in hyperspectral imagery: an overview. Image Processing, Applications and Systems Conference. Sfax, pp. 105–13 (5–7 Nov 2014)

[6] Geng, X., Sun, K., Ji, L., & Zhao, Y.: A high-order statistical tensor-based algorithm for anomaly detection in hyperspectral imagery. Scientific reports, 4, 6869 (2014)

[7] Chang, C. I.: Orthogonal subspace projection (OSP) revisited: a comprehensive study and analysis. IEEE Trans. Geosci. Remote Sens., 43(3), 502-518(2005)

[8] Johnson, R. J., Williams, J. P., & Bauer, K. W.: AutoGAD: An improved ICA-based hyperspectral anomaly detection algorithm. IEEE Transactions on Geoscience and Remote Sensing, 51(6), 3492-3503(2013)

[9] Geng, X., Ji, L., Zhao, Y., & Wang, F.: A small target detection method for the hyperspectral image based on higher order singular value decomposition (HOSVD). IEEE Geoscience and Remote Sensing Letters, 10(6), 1305-1308 (2013)

[10] Bourennane, S., & Fossati, C.: Improvement of small target detection based on tensorial filtering. Remote Sensing Letters, 6(10), 765-774. (2015)

[11] Lin, T., Marot, J., & Bourennane, S.: Small target detection improvement in hyperspectral image. In International Conference on Advanced Concepts for Intelligent Vision Systems (pp. 460-469). Springer, Cham (2013, October).

[12] Du, B., Zhao, R., Zhang, L., & Zhang, L.: A spectral-spatial based local summation anomaly detection method for hyperspectral images. Signal Processing, 124, 115-131 (2016)

[13] E. A. Ashton, “Detection of subpixel anomalies in multispectral infrared imagery using an adaptive Bayesian classifier,” IEEE Trans. Geosci. Remote Sens. 36, 506–517 (1998).

[14] Msellmi, B., Rabah, Z. B., & Farah, I. R.: A GRAPH BASED MODEL FOR SUB-PIXEL OBJECTS RECOGNITION. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium (pp. 7070-7073). IEEE. (2018, July)

[15] Kerekes, J. P.: Hyperspectral remote sensing subpixel object detection performance. (2011)

[16] Wu, Y., López, S., Zhang, B., Qiao, F., & Gao, L.: Approximate computing for onboard anomaly detection from hyperspectral images. Journal of Real-Time Image Processing, 1-16. (2018)

[17] Chang, C. I., Jiao, X., Wu, C. C., Du, Y., & Chang, M. L.: A review of unsupervised spectral target analysis for hyperspectral imagery. EURASIP Journal on Advances in Signal Processing, 2010(1), 503752. (2010)

[18] Wang, Z., Yin, Q., Li, H., & Hu, B.: Surface ship target detection in hyperspectral images based on improved variance minimum algorithm. In Eighth International Conference on Digital Image Processing (ICDIP 2016) (Vol. 10033, p. 100330R). International Society for Optics and Photonics. (2016, August)

[19] Yan, L., Noro, N., Takara, Y., Ando, F., & Yamaguchi, M.: Using hyperspectral image enhancement method for small size object detection on the sea surface. In Image and Signal Processing for Remote Sensing XXI (Vol. 9643, p. 96430H). International Society for Optics and Photonics (2015, October).

[20] Sumimoto, T., Kuramoto, K., Okada, S., Miyauchi, H., Imade, M., Yamamoto, H., T. Kunishi, H., "Machine vision for detection of the rescue target in the marine casualty," Proc. 20th International Conference on Industrial Electronics, Control and Instrumentation 1994 (IECON '94), vol.2, 723-726 (1994).

[21] Westall, P., Ford, J. J., O'Shea, P., Hrabar, S., "Evaluation of Maritime Vision Techniques for Aerial Search of Humans in Maritime Environments," Digital Image Computing: Techniques and Applications (DICTA), 176-183(2008).

[22] Doherty, P. and Rudol, P., "A UAV Search and Rescue Scenario with Human Body Detection and Geolocalization," AI 2007: Advances in Artificial Intelligence, Proc. 20th Australian Joint Conference on Artificial Intelligence, 1-13(2007).

[23] Mitsui, M., Murakami, Y., Obi, T., Yamaguchi, M., & Ohyama, N., "Color enhancement in multispectral image using the Karhunen-Loeve transform," Optical review, 12(2), 69-75, (2005).

[24] Hashimoto, N., Murakami, Y., Bautista, P. A., Yamaguchi, M., Obi, T., Ohyama, N., and Kosugi, Y., "Multispectral image enhancement for effective visualization," Optics express, 19(10): 9315-9329 (2011).

[25] Nascimento, J., Bioucas Dias, J., "Vertex component analysis: a fast algorithm to unmix hyperspectral data," Geoscience and Remote Sensing, IEEE Transactions on, 43(4), 898-910, (2005).

[26] Chang C. I., "Spectral information divergence for hyperspectral image analysis," Geoscience and Remote Sensing Symposium, 1999. IGARSS'99 Proc. IEEE, 1, 509-511 (1999).

[27] Keshava, N., and J. F. Mustard. "Spectral unmixing," Signal Processing Magazine, IEEE 19(1), 44-57(2002).

[28] Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B.: 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, 90(3), 337-352 (2004)

[29] Van der Meer, F. D., Van der Werff, H. M., Van Ruitenbeek, F. J., Hecker, C. A., Bakker, W. H., Noomen, M. F., ... & Woldai, T.: Multi-and hyperspectral geologic remote sensing: A review. International Journal of Applied Earth Observation and Geoinformation, 14(1), 112-128 (2012)

[30] Dalponte, M., Bruzzone, L., & Gianelle, D.: Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas. IEEE Transactions on Geoscience and Remote Sensing, 46(5), 1416-1427 (2008)

[31] Gono, K., Obi, T., Yamaguchi, M., Oyama, N., Machida, H., Sano, Y., ... & Endo, T.: Appearance of enhanced tissue features in narrow-band endoscopic imaging. Journal of biomedical optics, 9(3), 568-578 (2004)

[32] Manolakis, D., Marden, D., & Shaw, G. A.: Hyperspectral image processing for automatic target detection applications. Lincoln laboratory journal, 14(1), 79-116 (2003)

[33] Murakami, Y., Nomura, J., Ohyama, M., & Yamaguchi, M.: Fidelity evaluation of metallic luster in six-band high-dynamic-range imaging. Optical review, 19(3), 142-149 (2012)

[34] Banterle, F., Artusi, A., Debattista, K., & Chalmers, A.: Advanced High Dynamic Range Imaging: Theory and Practice. ISBN: 978-156881-719-4, AK Peters

[35] Brauers, J., Schulte, N., Bell, A. A., & Aach, T.: Multispectral high dynamic range imaging. In Color Imaging XIII: Processing, Hardcopy, and Applications (Vol. 6807, p. 680704). International Society for Optics and Photonics (2008)

[36] Hill, B., & Vorhagen, F. W.: U.S. Patent No. 5,319,472. Washington, DC: U.S. Patent and Trademark Office (1994)

[37] Lapray, P. J., Wang, X., Thomas, J. B., & Gouton, P.: Multispectral filter arrays: Recent advances and practical implementation. Sensors, 14(11), 21626-21659 (2014)

[38] Gupta, R., & Hartley, R. I.: Linear pushbroom cameras. IEEE Transactions on pattern analysis and machine intelligence, 19(9), 963-975 (1997)

[39] Abel, J. S., & Smith, J. O.: Restoring a clipped signal. In [Proceedings] ICASSP 91: 1991 International Conference on Acoustics, Speech, and Signal Processing (pp. 1745-1748). IEEE (1991)

[40] Dahimene, A., Noureddine, M., & Azrar, A.: A simple algorithm for the restoration of clipped speech signal. Informatica, 32(2) (2008)

[41] Zhang, X., & Brainard, D. H.: Estimation of saturated pixel values in digital color imaging. JOSA A, 21(12), 2301-2310 (2004)

[42] Guo, D., Cheng, Y., Zhuo, S., & Sim, T.: Correcting over-exposure in photographs. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (pp. 515-521). IEEE (2010)

[43] Zhang, H., He, W., Zhang, L., Shen, H., & Yuan, Q.: Hyperspectral image restoration using low-rank matrix recovery. IEEE Transactions on Geoscience and Remote Sensing, 52(8), 4729-4743. (2013)

[44] Liao, W., Goossens, B., Aelterman, J., Luong, H. Q., Pižurica, A., Wouters, N., ... & Philips, W.: Hyperspectral image deblurring with PCA and total

variation. In 2013 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) (pp. 1-4). IEEE. (2013)

[45] Tzeng, D. Y., & Berns, R. S.: A review of principal component analysis and its applications to color technology. Color Research & Application: Endorsed by Inter‐Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur, 30(2), 84-98 (2005)

[46] Takara, Y., Manago, N., Saito, H., Mabuchi, Y., Kondoh, A., Fujimori, T., ... & Kuze, H.: Remote sensing applications with NH hyperspectral portable video camera. In Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications IV (Vol. 8527, p. 85271G). International Society for Optics and Photonics. (2012)

[47] Yamaguchi, M., Haneishi, H., & Ohyama, N.: Beyond red–green–blue (RGB): spectrum-based color imaging technology. Journal of Imaging Science and Technology, 52(1), 10201-1 (2008)

[48] Haneishi, H., Hasegawa, T., Hosoi, A., Yokoyama, Y., Tsumura, N., & Miyake, Y.: System design for accurately estimating the spectral reflectance of art paintings. Applied Optics, 39(35), 6621-6632 (2000)

[49] Wu, D., & Sun, D. W.: Colour measurements by computer vision for food quality control–A review. Trends in Food Science & Technology, 29(1), 5-20 (2013)

[50] Abe, T., Murakami, Y., Yamaguchi, M., Ohyama, N., & Yagi, Y.: Color correction of pathological images based on dye amount quantification. Optical review, 12(4), 293-300 (2005)

[51] Tominaga, S., & Wandell, B. A.: Standard surface-reflectance model and illuminant estimation. JOSA A, 6(4), 576-584 (1989)

[52] Abrams, M. J., Ashley, R. P., Rowan, L. C., Goetz, A. F., & Kahle, A. B.: Mapping of hydrothermal alteration in the Cuprite mining district, Nevada, using aircraft scanner images for the spectral region 0.46 to 2.36 μm. Geology, 5(12), 713-718 (1977)

[53] Zhang, C., & Kovacs, J. M.: The application of small unmanned aerial systems for precision agriculture: a review. Precision agriculture, 13(6), 693-712 (2012)

[54] Liang, H.: Advances in multispectral and hyperspectral imaging for archaeology and art conservation. Applied Physics A, 106(2), 309-323 (2012)

[55] Edelman, G. J., Gaston, E., Van Leeuwen, T. G., Cullen, P. J., & Aalders, M. C. G.: Hyperspectral imaging for non-contact analysis of forensic traces. Forensic science international, 223(1-3), 28-39 (2012)

[56] Lu, G., & Fei, B.: Medical hyperspectral imaging: a review. Journal of biomedical optics, 19(1), 010901 (2014)

[57] Nasrabadi, N. M.: Hyperspectral target detection: An overview of current and future challenges. IEEE Signal Processing Magazine, 31(1), 34-44 (2014)

[58] Melgani, F., & Bruzzone, L.: Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on geoscience and remote sensing, 42(8), 1778-1790 (2004)

[59] Mountrakis, G., Im, J., & Ogole, C.: Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 66(3), 247-259 (2011)

[60] Belgiu, M., & Drăguţ, L.: Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31 (2016)

[61] Zhang, L., Zhang, L., & Du, B.: Deep learning for remote sensing data: A technical tutorial on the state of the art. IEEE Geoscience and Remote Sensing Magazine, 4(2), 22-40 (2016)

[62] Pan, B., Shi, Z., & Xu, X.: MugNet: deep learning for hyperspectral image classification using limited samples. ISPRS Journal of Photogrammetry and Remote Sensing (2017)

[63] Pan, B., Shi, Z., Zhang, N., & Xie, S.: Hyperspectral image classification based on nonlinear spectral–spatial network. IEEE Geoscience and Remote Sensing Letters, 13(12), 1782-1786 (2016)

[64] Hu, W., Huang, Y., Wei, L., Zhang, F., & Li, H.: Deep convolutional neural networks for hyperspectral image classification. Journal of Sensors (2015)

[65] Petersson, H., Gustafsson, D., & Bergstrom, D.: Hyperspectral image analysis using deep learning—a review. In Image Processing Theory Tools and Applications (IPTA), 2016 6th International Conference on 1-6. IEEE. (2016)

[66] Makantasis, K., Karantzalos, K., Doulamis, A., & Doulamis, N.: Deep supervised learning for hyperspectral data classification through convolutional neural networks. In Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International 4959-4962 (2015)

[67] Ruder, S.: An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747 (2016)

[68] Ma, X., Wang, H., & Geng, J.: Spectral–spatial classification of hyperspectral image based on deep auto-encoder. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(9), 4073-4085 (2016)

[69] Manolakis, D., & Shaw, G.: Detection algorithms for hyperspectral imaging applications. IEEE signal processing magazine, 19(1), 29-43 (2002)

[70] Li, W., Wu, G., & Du, Q. Transferred Deep Learning for Anomaly Detection in Hyperspectral Imagery. IEEE Geosci. Remote Sensing Letters, 14(5), 597-601 (2017)

[71] Murray-Krezan, J., Neumann, J. G., & Leathers, R. A.: Small object hyperspectral detection from a low-flying UAV. In Signal and Data Processing of Small Targets 2008 (Vol. 6969, p. 69691C). International Society for Optics and Photonics. (2008)

[72] François, C., and others: GitHub Repository, 2015. https://keras.io. Accessed 1 November 2018

[73] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Kudlur, M.: Tensorflow: a system for large-scale machine learning. In OSDI (Vol. 16, 265-283) (2016)

[74] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J.: Scikit-learn: Machine learning in Python. Journal of machine learning research, 12(Oct), 2825-2830 (2011)

[75] Subhajit Chaudhury,: GitHub Repository, 2016. https://github.com/subhajitchaudhury/deephypercnn. Accessed 1 November 2018

List of Publications

Journal:

[1] Yan, L., Yamaguchi, M., Noro, N., Takara, Y., & Ando, F.: A novel two-stage deep learning-based small-object detection using hyperspectral images. Optical Review, 26(6), 597-606. (2019) https://doi.org/10.1007/s10043-019-00528-0

[2] Yan, L., Yamaguchi, M., Noro, N., Takara, Y., & Ando, F.: Effect of the restoration of saturated signals in hyperspectral image analysis and color reproduction. Optical Review, https://doi.org/10.1007/s10043-020-00630-8

International Conference:

[3] Yan, L., Noro, N., Takara, Y., Ando, F., & Yamaguchi, M.: Using hyperspectral image enhancement method for small size object detection on the sea surface, Proc. SPIE 9643, Image and Signal Processing for Remote Sensing XXI, 96430H (15 October 2015); https://doi.org/10.1117/12.2194606

Domestic Conference:

[4] Yan, L., Yamaguchi, M., Noro, N., Takara, Y., & Ando, F.: Human detection method using airborne hyperspectral images, Industrial applications of

Hyperspectral and Multispectral technologies, ハイパースペクトル・マルチスペクトルデータの計測と産業応用 研究会 講演予稿集, p. 6, Dec. 2018.

[5] Yan, L., Yamaguchi, M., Noro, N., Takara, Y., & Ando, F.: Effect of the restoration of saturated signal in hyperspectral image analysis and color reproduction, Visual/Media Computing Conference 2019, 第 47 回画像電子学会年次大会, Jun. 2019.