

1

Overview and Introduction

The past few years have witnessed tremendous advances in hyperspectral imaging where statistical signal processing has played a pivotal role in driving algorithm design and development for hyperspectral data exploitation. It has attracted attention of those who come from different disciplinary areas by exploring new applications and making connections between remote sensing and other engineering fields. In recent years, there has been a significant increase in participation in various conferences and venues related to this area, which in turn has provided evidence that hyperspectral signal and image processing has broken away from traditional spatial domain analysis-based remote sensing and successfully branched out to stand alone as a single research topic, similar to signal processing that evolved as a separate area from communications in the late 1970s. On the contrary issues related to high spectral resolution provided by hyperspectral imaging sensors have also changed the ways in which algorithms are designed and developed. As a consequence, many problems such as subpixels and mixed pixels that are generally encountered in hyperspectral image processing, but do not occur in classical two-dimensional (2D) image processing, have become major issues for traditional spatial domain-based techniques. This is because the concept of “*seeing-is-believing*” by visual inspection, which has been widely used in image processing, cannot resolve issues of targets that are completely embedded in a single pixel or partially but do not fully occupy a single pixel, in which case only spectral properties can be used to characterize such targets for data analysis. Therefore, to distinguish such spectral characterization-based analysis from the traditional spatial domain-based analysis, the former is referred to as *nonliteral* analysis as opposed to the latter termed as *literal* analysis.

Due to complicated environments in real-world problems many uncontrollable parameters are also beyond our grip. In order to explore insights into algorithm design, the use of synthetic images to simulate various scenarios to substantiate designed algorithms for performance analysis becomes an effective proof-of-concept evaluation tool. Such synthetic images can be simulated by either real image scenes or laboratory data sets for various applications. Unfortunately, such synthetic image-based computer simulations have received little attention in the past. Accordingly, one of the major features that readers will find in this book is an extensive use of synthetic image-based experiments in algorithm design and analysis for qualitative as well as quantitative performance evaluation. Another unique feature of this book is that the algorithms derived and developed in this book can be implemented with little difficulty via the MATLAB, a widely accepted software package developed by the MATHWORK for engineering applications. This advantage allows readers to implement their algorithms. To further facilitate this benefit MATLAB codes of many

popular algorithms developed in this book are also made available in the appendix at the end of this book. Most importantly, this book has also expanded its use of images in Chang (2003a) to include two more popular image scenes, Purdue's Indian Pine test site in Indiana and Cuprite image scene in Nevada, both of which are available on web site so that they can be used by those who develop new algorithms, to validate and evaluate their designed algorithm for performance analysis as well as to conduct their own comparative study. Last but not the least, this book includes an appendix that compiles many algorithms developed in the Remote Sensing Signal and Image Processing Laboratory (RSSIPL) at the University of Maryland, Baltimore County (UMBC). Such an algorithm compendium should serve as a valuable guide for people who are interested in applications of hyperspectral data processing.

1.1 Overview

Hyperspectral signal and image processing has become a fast-growing area that bridges communities of remote sensing and signal/image processing due to the fact that many problems arising in the former can be reformatted and solved by the latter. A good example is an increasing number of conferences in both communities and a wide range of publications in both signal and image processing journals and traditional remote sensing journals. Particularly, in recent years significant research and development in hyperspectral imaging has resulted in at least hundreds, if not thousands, of publications in various journals. Many new findings have been reported in many annual meetings and venues such as IEEE International Geoscience and Remote Sensing Symposia; SPIE conferences on Defense and Security (previously known as AeroSense); Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery (annually held in April); SPIE International Symposium on Optical Science & Technology (*Remote Sensing Symposium*, specifically Conferences on *Satellite Data Compression, Communication, and Processing* and Conferences on *Imaging Spectrometry*, annually held in August); Conference on Imaging Spectrometry; EOS/SPIE Symposium on Remote Sensing; IEEE GRSS Workshop on Hyperspectral Image and Signal Processing—Evolution in Remote Sensing (WHISPERS); and so forth. Accordingly, any attempt to cover in a single book all possible areas in this field would be impossible and unrealistic. Keeping this reality in mind the book is written based on personal preference and is only focused on recent works that have been done mostly in RSSIPL at UMBC, but have not been covered in my previous book (Chang, 2003a). Therefore, there is not much overlap between this book and Chang (2003a). Hence this book can be considered as a sequel of Chang (2003a). Specifically, this book takes a rather different yet unique approach compared with that adopted in Chang (2003a), by treating hyperspectral image processing and hyperspectral signal processing as two separate subjects where the former processes a hyperspectral image as an image cube and the latter considers a hyperspectral signature as a one-dimensional signal so that no sample correlation such as spectral correlation among pixels in a hyperspectral image cube can be taken into account and used for algorithm design. Within this context the topics presented in this book are organized in the following order.

- I. Preliminaries
 - A. Hyperspectral Image Processing
- II. Endmember Extraction
- III. Supervised Linear Hyperspectral Mixture Analysis
- IV. Unsupervised Hyperspectral Image Analysis
- V. Hyperspectral Information Compression
- B. Hyperspectral Signal Processing

VI. Hyperspectral Signal Coding

VII. Hyperspectral Signal Characterization

C. Applications

Appendix: Algorithm Compendium

Several recent books on hyperspectral imaging are also available in the public domain (Chang, 2003a; 2006b; 2007a; Plaza and Chang, 2007a; Varshney and Arora, 2004) and the subjects covered in these books are somewhat selective. For example, the book by Varshney and Arora (2004) is focused on some specific techniques, for example, independent component analysis, support vector machines, and Markov random field. The book by Chang (2003a) is primarily devoted to nonliteral statistical signal processing techniques developed for subpixel detection and mixed pixel classification. The two books edited by Chang (2006b) and (2007a) are collections of most recent results contributed by researchers who are experts and currently active in hyperspectral imaging. Another book edited by Plaza and Chang (2007a) intends to target specific topics in high-performance computing, which has grown rapidly due to the need of processing enormous data volumes and has also become increasingly important in remote sensing data processing. Unlike the above-mentioned books this book explores many interesting research topics resulting from issues that are generally either overlooked or neglected in multispectral imagery as well as issues that were not addressed or fully explored in Chang (2003a).

1.2 Issues of Multispectral and Hyperspectral Imageries

Because of its low spectral resolution a multispectral image pixel vector usually does not have information as rich as a hyperspectral image pixel vector does. In this case, multispectral image processing must rely on image spatial information and correlation to make up insufficient spectral information resulting from a few discrete spectral bands. Therefore, an early development of multispectral image processing has focused on spatial domain-based techniques. However, with recent advent of very high-spectral resolution hyperspectral imaging sensors many material substances that cannot be resolved by multispectral imaging sensors can now be uncovered by hyperspectral imagers for data analysis. As a consequence, targets or objects of interest for multispectral and hyperspectral image analyses are quite different. In multispectral image analysis land cover/land use is often of major interest. Therefore, the developed techniques generally perform pattern classification and analysis in the sense that every single pixel of an image must be classified into one of a number of pattern classes, each of which corresponds to one particular spatial class. On the contrary, the objects of interest in hyperspectral image analysis are usually targets with particular spectral characteristics such as man-made targets, anomalies, or rare targets. The targets of these types generally appear either in a form mixed with a number of material substances or at subpixel level with targets embedded in a single pixel vector due to their size that is smaller than the ground sampling distance (GSD). Besides, these types of targets usually appear unexpectedly and their probabilities of occurrence are also low. Most importantly, their sample pool may also be relatively small and their sizes may only have limited spatial extent. As a consequence, such targets may not be easy to be visually identified or inspected with prior knowledge; thus, they can be considered as insignificant targets but are indeed of major interest from an intelligence or information point of view. For example, these targets may include special spices in agriculture and ecology, toxic wastes in environmental monitoring, rare minerals in geology, drug/smuggler trafficking in law enforcement, military vehicles in combat, abnormality in battlefields, landmines in war zones, chemical/biological agents in bioterrorism, weapon concealment and mass graves in intelligence gathering, and so on. Under such circumstances, they can be only detected at mixed or subpixel

level and traditional spatial domain (i.e., literal)-based image processing techniques may not be suitable or effective. So, the extraction of such targets must rely on their spectral profiles and the techniques developed for hyperspectral image analysis should perform *target*-based detection, discrimination, classification, identification, recognition, and quantification as opposed to *pattern*-based multispectral imaging techniques. As a result, a direct extension of multispectral imaging techniques to hyperspectral imagery may not be effective in hyperspectral data exploitation because pattern class information and correlation provided by these targets may be too little to be used for performing hyperspectral image analysis. In order to address this issue the techniques in Chang (2003a) are developed directly from a hyperspectral imagery point of view for spectral detection and classification. This book expands the scope of Chang (2003a) to cover a wider range of applications including endmember extraction, unsupervised target detection, information compression, and hyperspectral signal coding and characterization, none of which is studied in Chang (2003a).

1.3 Divergence of Hyperspectral Imagery from Multispectral Imagery

The hyperspectral imagery has changed the way we think of multispectral imagery. This is because we now have hundreds of contiguous spectral bands available at our disposal. So, one major issue is how to effectively use and take advantage of spectral information provided by these hundreds of spectral bands for various applications in data exploitation, for example, target detection, discrimination, classification, quantification, and identification. This interesting issue can be addressed by the following two interesting examples. The first example uses real-to-complex analysis to illustrate why it is inappropriate to simply extend multispectral imaging techniques to process hyperspectral imagery. The second example uses the well-known pigeon-hole principle in discrete mathematics (Epp, 1995) to illustrate how hyperspectral imagery can be addressed by a rationale completely different from that used for multispectral imagery.

1.3.1 *Misconception: Hyperspectral Imaging is a Natural Extension of Multispectral Imaging*

While dealing with hyperspectral imagery there is a general consensus that hyperspectral imagery is a natural extension of multispectral imagery based on an assumption that a hyperspectral image has more spectral bands for data collection than a multispectral image does. As a result, it may lead to a misconception that hyperspectral imaging problems can be solved by multispectral imaging techniques by simply taking advantage of its expanded spectral bands. A similar misconception also occurs in hyperspectral data compression where researchers in data compression community consider a hyperspectral data as an image cube so that 3D image compression processing techniques developed for videos can be simply applied to hyperspectral imagery as a 3D image cube without extra precaution (see Part V: Chapters 19–23). Unfortunately, over the past few years these misconceptions have somewhat directed the way we design and develop hyperspectral imaging techniques.

To understand the fundamental difference between multispectral imaging and hyperspectral imaging, we use a simple mathematical example to illustrate a similar misconception, which is finding derivatives in real analysis and complex analysis. Since real variables can be considered as real parts of complex variables, this may lead to a brief that real analysis is a special case of complex analysis, which is certainly not true. One piece of clear evidence is derivatives. When a derivative is calculated in the real line, the direction with respect to which a derivative is calculated along the real axis is constrained either to the left or to the right. However, the direction along

which a derivative is calculated by complex analysis can be along any curve in the complex plane. As a result, calculating a complex derivative is more sophisticated than simply extending the way derivatives are calculated in real analysis. A natural extension of real derivatives is partial derivatives in complex analysis along two axes: x -axis and y -axis. However, it is not true for any derivative calculated in the complex plane. This is because the direction along which the derivative is calculated is not only limited to x - and y -axes but it must also take into account all directions that are more likely curves instead of lines. When such a derivative occurs it is called total differentiable or analytic and must satisfy the so-called Cauchy–Riemann equation that allows a differentiable complex variable to be expanded as a power series which is much stronger than only derivatives. This simple example explains why complex analysis is not a natural extension of real analysis and a direct extension of real derivatives to complex derivatives as partial derivatives can only achieve limited success to some extent. This example sheds some light on a similar key difference between multispectral and hyperspectral images. In its early days multispectral imagery has been used in remote sensing mainly for land cover/land use classification in agriculture, disaster assessment and management, ecology, environmental monitoring, geology, geographical information system (GIS), and so on. In these applications, low spectral resolution multispectral imagery may provide sufficient information for data analysis and the techniques developed for multispectral image processing are primarily based on pattern classes that take advantage of spatial correlation to perform various tasks. Compared to multispectral imagery, hyperspectral imagery utilizes hundreds of contiguous spectral bands to perform target-class analysis. This is the major difference between hyperspectral imagery and multispectral imagery. Specifically, the objects of interest in hyperspectral imagery are no longer patterns of large areas as considered in multispectral imagery. Instead, hyperspectral image analysts are interested in those objects that cannot be visualized by inspection or with prior knowledge due to limited extent of their spatial presence. As a result, hyperspectral imaging is generally developed to perform target class–based image analysis where image background is usually of no interest. Such examples include anomaly detection, endmember extraction, man-made target detection, and so on, where the spatial information provided by these objects of interest is generally very little. So, if the hyperspectral imagery is treated as a natural extension of the multispectral imagery, its success can be very limited due to its use of spatial information to perform *pattern class–based image analysis* rather than *target class–based image analysis*, a similar dilemma that also occurs between real and complex derivatives. Accordingly, we must reinvent the wheel and re-design and develop new hyperspectral imaging techniques rather than directly derive those adopted from multispectral image techniques. One promising approach is the use of the following pigeon-hole principle described in the following section.

1.3.2 Pigeon-Hole Principle: Natural Interpretation of Hyperspectral Imaging

Suppose that there are p pigeons flying into L pigeon holes (nests) with $L < p$. According to the pigeon-hole principle, there exists at least one pigeon hole that must accommodate at least two or more pigeons. Now, assume that L is the total number of spectral bands and p is the number of targets of interest. By virtue of the pigeon-hole principle, we can interpret a pigeon hole as a spectral band while a pigeon is considered as a target (or an object) of interest. With this interpretation if $L > p$, a spectral band can be used to detect, discriminate, and classify a distinct target. Since there are hundreds of spectral bands available from hyperspectral imagery, technically speaking, hundreds of spectrally distinct targets can be accommodated by these spectral bands, namely one target by one particular spectral band. In order to materialize this idea, three issues need to be addressed. First, the number of spectral bands, L , must be greater than or equal to the number of targets of interest, p , that is, $L \geq p$. This seems always true for hyperspectral imagery, but is not

necessarily valid for multispectral imagery, where $L < p$ in the latter is usually true. For example, 3-band SPOT multispectral data may have difficulty with classifying more than three target substances present in the data using the pigeon-hole principle. However, the benefit of $L \geq p$ also gives rise to a challenging issue known as “curse of dimensionality” (Duda and Hart, 1973), that is, “what is the true value of p if $L \geq p$.” This has been a long-standing issue for any hyperspectral image analyst to resolve because it is nearly impossible to know the exact value of p in real-world problems. Moreover, even if the value of p can be provided by prior knowledge it may not be reliable due to many unexpected factors that cannot be known *a priori*. In multivariate data analysis, the value of p is described by the so-called intrinsic dimensionality (ID) (Fukunaga, 1990), which is defined as the minimum number of parameters used to specify the data. However, this concept is only of theoretical interest. No method has been proposed regarding how to find it in the literature. A common strategy is to estimate the p on a trial-and-error basis. A similar problem is also encountered in passive array processing where the number of signal sources, p , arriving at a linear array of sensors is of major interest. In order to estimate this number, two criteria, an information criterion (AIC) suggested by Akaike (1974) and minimum description length developed by Schwarz (1978) and Rissanen (1978), have been widely used to estimate the value of p . Unfortunately, a key assumption made on these criteria is that the noise must be independent and identically distributed, a fact that is usually not a valid assumption in hyperspectral images as shown in Chang (2003a) and Chang and Du (2004). In order to cope with this dilemma, a new concept called virtual dimensionality (VD) was coined and suggested by Chang (2003a) to estimate the number of spectrally distinct signatures in hyperspectral imagery. It is also based on the pigeon-hole principle where VD is used to estimate the number of pigeons with the total number of spectral bands interpreted as the number of pigeon holes. The last issue to be addressed is that once a spectral band is being used to accommodate one target, it cannot be used again to accommodate another distinct target. One way to do so is to perform orthogonal subspace projection (OSP) developed by Harsanyi and Chang (1994) on a space linearly spanned by the already found targets to find an orthogonal complement space from which only new targets can be generated. Equivalently speaking, the spectral bands used to accommodate previous targets cannot be used again to accommodate a new target. Through a series of such OSP operations no two distinct targets will be specified and accommodated by a single spectral band. In other words, all the found targets must be accommodated in separate mutual orthogonal subspaces. In terms of the pigeon-hole principle it implies that no two pigeons will be allowed to fly into a single pigeon hole. Here, one remark is noteworthy. When it says that one target is accommodated and specified by one spectral band, it simply means that the target can be best spectrally characterized by this particular band compared to other bands. So, this band is chosen to be its identity like its fingerprint or DNA. If two targets happen to have the same band being used for their best spectral characterization then there is no way to discriminate one from the other. In this case, it implies that two pigeons fly into the same pigeon hole. More specifically, one pigeon hole is used to accommodate two flying-in pigeons, both of which reside in a single pigeon hole.

Once these three issues, that is, (1) $L \geq p$, (2) determination of p , and (3) no two distinct target signatures to be accommodated by a single spectral band, are resolved, the idea of applying the pigeon-hole principle to hyperspectral data exploitation can be realized and becomes feasible. More specifically, using spectral bands as a means to perform detection, discrimination, classification, and identification without accounting for spatial information or correlation provides an alternative approach, to be called nonliteral analysis as opposed to the spatial domain-based approach, to be called literal analysis. Such a nonliteral analysis is particularly important for two types of targets. One is that targets are small or insignificant due to their limited spatial presence and cannot be effectively captured by spatial correlation or information. The other is that targets of the same

type are spatially separated so that their spatial correlation is actually very weak and little in which case the spatial domain-based literal analysis may have difficulty in finding them spatially correlated. The only way to group them together is based on their spectral characteristics regardless of where they are spatially located.

Interestingly, the pigeon-hole principle also sheds light on differentiation of hyperspectral imagery from multispectral imagery. Through the relationship between the total number of spectral bands, L , and the number of signal sources to be accommodated, p , discussed above, a multiple-band remote sensing image can be considered as a hyperspectral image if $L \geq p$ and a multispectral image otherwise (i.e., $L < p$). More details of this interpretation can be found in Chapter 31.

Furthermore, VD can be also interpreted by the pigeon-hole principle, and its potential in hyperspectral data exploitation has been demonstrated in many applications, for example, linear spectral mixture analysis (Chang, 2006c), dimensionality reduction (Wang and Chang, 2006a, 2006b), band selection (Chang and Wang, 2006), and so on. Chapter 5 will revisit VD for more details.

1.4 Scope of This Book

While writing this book it is important to consider hyperspectral image processing and hyperspectral signal processing as two different research areas and treat them separately. When hyperspectral data are processed as image cubes, it is called hyperspectral image processing where data samples are image pixel vectors and both spectral and spatial correlation among image pixel vectors can be made available for data processing. On the other hand, when hyperspectral data are processed as signatures it is called hyperspectral signal processing where a signature is a one-dimensional signal, which represents its spectral profile over a range of wavelengths for signature characterization. In this case, only interband spectral correlation within the signature is available for data processing and no other information such as sample spatial or spectral correlation used in hyperspectral image processing is available for signature processing. Such hyperspectral signals include data obtained from laboratories, databases, and spectral libraries where no data sample spatial/spectral correlation is available. Therefore, techniques developed for hyperspectral image processing may not be directly applicable to hyperspectral signal processing and vice versa. Unfortunately, it seems that there is no concern in distinguishing one from another when it comes to algorithm design. This book is believed to be the first to do so by treating hyperspectral image processing and hyperspectral signal processing in two separate categories: Category A: Hyperspectral Image Processing treated in Parts II–V; Category B: Hyperspectral Signal Processing treated in Parts VI–VII.

In order to make this book self-contained, preliminaries are also included as Part I to cover basic knowledge that provides readers with necessary background required to read this book. In particular, it integrates many scattering results into different chapters so that readers can follow through materials easily without looking for other references. Each chapter in Part I can be read independently with little interruption while also keeping the flow and all the chapters coherent each other.

Part II is endmember extraction that is one of most crucial tasks in hyperspectral data exploitation and has recently become increasingly important due to significantly improved high spatial and spectral resolution provided by hyperspectral imaging sensors. According to the definition given by Schowengerdt (1997), an endmember is an idealized, pure signature for a class, more specifically, spectral class. For multispectral imagery, an endmember may nowhere be found since most data sample vectors may be heavily mixed due to low spatial and spectral resolution. As a result, the importance of endmember extraction has been overlooked and not been a major subject in multispectral image analysis. By contrast, with recent advances of hyperspectral imaging sensors many subtle material substances that cannot be resolved by multispectral imagery can be now

revealed by hyperspectral imagery. These substances are generally not known *a priori* and can be only diagnosed by high spectral resolution. Endmembers are considered to be one of such substances. In general, their existence in image data cannot be detected visually. Most importantly, once endmembers are present, their spatial extent is relatively limited. Besides, their sample pools are also very small. Accordingly, they may appear as anomalies. In this case, spatial characteristics offer little advantage in finding endmembers. In the past, the image classification in multispectral image processing has been often performed by pattern classification (land use/land cover classification) where each image pixel must be classified into a particular class in accordance with a certain classification criterion. However, endmembers are generally rare. Unless they are treated and extracted as targets of interest, their detection and extraction is very challenging. Additionally, because of the lack of spatial patterns specified by endmembers the effectiveness of endmember detection or extraction will be very likely to be compromised by spatial-based pattern classification techniques. In order to address this issue, Chang (2003a) has focused on target classification than on pattern classification, in which case only targets of interest are of major concern where image background is only used for suppression. However, such an important issue of endmember extraction was not investigated and explored in Chang (2003a), when this subject was not mature but now will be one of the major subjects in this book studied in great detail in Part II.

Part III revisits supervised linear spectral mixture analysis (SLSMA), which was discussed in great length in Chang (2003a). This part rederives a least squares-based orthogonal subspace projection (LSOSP) from the signal-to-noise ratio (SNR)-based orthogonal subspace projection so that LSOSP and OSP essentially operate the same matched filter subject to a constant κ , which accounts for least squares estimation error. More specifically, LSOSP performs as an estimator by including the κ , while OSP operates as a detector by setting $\kappa = 1$. By using different matched signatures LSOSP and OSP can interpret many commonly used operators such as constrained energy minimization (CEM) in Chang (2003a) and RX detector developed by Reed and Yu (1990). Furthermore, OSP and LSOSP can be extended in three different directions. One is to replace the least squares error criterion with Fisher's ratio to derive Fisher's LSMA (FLSMA). Another is to impose weight constraints on spectral bands to derive weighted abundance-constrained LSMA (WAC-LSMA). Finally, a third direction introduces a nonlinear kernel into LSMA to derive kernel-based LSMNA (KLSMA).

Part IV extends SLSMA developed in Part III to unsupervised LSMA (ULSMA) where prior knowledge of signature information is not available. Under such circumstance two major issues that do not occur in supervised analysis need to be addressed. One is how many signature sources of interest to be used for LSMA and the second is how to find them. Once these issues are resolved ULSMA becomes SLSMA where approaches presented in Part III are readily applied.

Part V is hyperspectral information compression. One challenging issue in processing hyperspectral imagery is its huge data volume, which may result in high computational cost of data processing, long delay of data transmission and communications, and difficult management of data storage and archiving. Another is how to compress spectral information resulting from highly correlated spectral bands without sacrificing vital information. The first issue can be addressed by developing techniques reducing data size, referred to as data reduction/compression, while the second issue can only be addressed by developing techniques removing redundant information, referred to as information compression. These two types of compression are completely different and should be dealt with separately. Unfortunately, many hyperspectral data compression techniques have not taken this distinction into account. But, it is important to differentiate information compression from data compression since the compression ratio used in data compression is measured by data size, which does not imply compression of information. In other words, information compression is determined by various applications with specific information required to be

retained during a compression process. This type of information compression can be considered as exploitation-based lossy compression. To address this issue, the commonly used terminology, hyperspectral data compression, is referred to as hyperspectral information compression in this book and can be interpreted as exploitation-based lossy hyperspectral compression, which includes two major spectral compression techniques, spectral dimensionality reduction and spectral band selection, each of which will be discussed in great detail in Part V.

Up to now the hyperspectral data considered in previous parts are image cubes where all the data sample vectors are image pixel vectors. However, in many situations the hyperspectral data may only be obtained as signature vectors by nonimage sensors or from spectral libraries or databases. In this case, the data to be dealt with is a one-dimensional hyperspectral signal as a signature vector rather than as a pixel vector in three-dimensional image cube. So, Category B in this book is primarily focused on hyperspectral signal processing, which consists of two parts, Part VI and Part VII. Part VI considers hyperspectral signal coding where information compression is performed on a hyperspectral signature vector to capture its unique spectral profile to serve its fingerprint for signature discrimination, detection, classification, and identification. In other words, instead of considering image data as a 3D image cube, the idea of hyperspectral signal coding is to explore spectral characteristics and further to capture changes in the spectral profile within a single signature vector as spectral marks so that a single signature vector can be encoded by its fingerprint as a code word to represent its identity. On the other hand, hyperspectral signal coding can also be considered as quantization that discretizes hundreds of spectral values into a finite set of discrete values. So, it can be viewed as an analog-to-digital (A/D) converter and intends to find the best possible representation of a hyperspectral signature vector for a given bit rate. For a multispectral signature vector the spectral resolution is low and only a few spectral values are available for quantization. So, signature coding may not be effective to characterize spectral signature properties. This may no longer be true for a hyperspectral signature vector where hundreds of contiguous spectral bands may provide sufficient information for spectral characterization. Interestingly, hyperspectral signature coding has never been of major interest in hyperspectral data analysis. This part investigates three types of hyperspectral signal coding: binary coding, vector coding, and progressive coding, where the binary coding can be viewed as memoryless coding as opposed to the vector coding and progressive coding, which can be regarded as memory coding. Comparing the hyperspectral signature coding in Part VI that makes hard decisions on the spectral profile of a signature vector, Part VII presents techniques that make soft decisions on a signature vector to perform hyperspectral signature analysis in the sense of hyperspectral signature characterization. In this case, the knowledge of a reference signature is generally required for a hyperspectral signature vector to be characterized. Unfortunately, hyperspectral signature analysis via spectral characterization has not received much attention in the last few years. Part VII investigates this issue by developing three different approaches: band selection for signature characterization, Kalman filter for signature estimation, and wavelets for signature representation.

The last category of this book is Category C: Applications, which show how hyperspectral imaging techniques can be applied to various problems such as size estimation of subpixel targets, concealed target detection, and how to take advantage of hyperspectral imaging techniques to resolve issues of multispectral imagery. Specifically, a new application of hyperspectral imaging to magnetic resonance imaging is included to demonstrate its utility in medical imaging.

To conclude this book, an appendix is also included for readers' reference. It is an algorithm compendium that compiles important algorithms developed in the RSSIPL at UMBC.

1.5 Book's Organization

This book is organized in accordance with the order laid out by seven parts in three categories presented in the previous section. Each part can be read independently while keeping sufficient correlation with other parts.

1.5.1 Part I: Preliminaries

The preliminaries in Part I help readers grasp sufficient knowledge to follow this book without difficulty. It consists of six chapters.

Chapter 2 is Fundamentals of Subsample and Mixed Sample Analyses. It uses a simple example to illustrate issues of subsamples and mixed samples encountered in detection and classification. It then walks through various approaches using hard and soft decisions for subsample detection and mixed sample classification. It includes many techniques currently being used and available in the literature.

Chapter 3 introduces Three-Dimensional Receiver Operating Characteristics (3D ROC) Analysis that can be used as an evaluation tool for soft decision-making performance for hyperspectral target detection and classification. An ROC curve is defined as a curve plotted based on detection probability versus false alarm probability. An analysis that uses ROC curves to evaluate the effectiveness of a Neyman–Pearson detector is called ROC analysis. A major advantage of ROC analysis is that there is no need of specifying a particular cost function. For example, least squares error or signal-to-noise ratio may be a good criterion for detection of problems in signal processing and communications, but may not be appropriate to measure image quality or classification accuracy. This is essentially true when it comes to design of computer-aided diagnostic systems where their effectiveness is measured by their end users in which case the cost function is generally human errors. Furthermore, ROC analysis is developed for detection in the context of binary hypothesis testing problems. In chemical/biological warfare (CBW) defense, estimation of chemical/biological (CB) agent abundance is more critical than CB agent detection since the lethal level of concentration of different CB agents poses different threats. The detection-based ROC curves cannot address this need. Chapter 3 is included to resolve this issue where a 3D ROC analysis is developed by creating a third dimension to specify target abundance so that a 3D-ROC curve can be generated and plotted based on three parameters, detection probability, P_D , false alarm probability, P_F , and threshold τ . Consequently, the traditional detection-based ROC curves, referred to as 2D ROC curves, become a special case of 3D ROC curves. As noted, most hyperspectral imaging techniques are actually derived from various aspects of estimation, which produce abundance fractions of signatures of interest such as linear spectral mixture analysis. In order to evaluate their performance for quantitative analysis the estimated abundance fractions must be converted to hard decisions via a threshold τ . The 3D ROC analysis provides a feasible tool for this purpose.

Chapter 4 is Design of Synthetic Image Experiments. One of major difficulties in algorithm design is how to evaluate various algorithms objectively and impartially on a fair common ground. In doing so, the first concern is the data to be used for experiments that must be available and assessable for those who are interested in comparing their designed algorithms to others. This can be done by using data sets in the public domain. A second concern is that the experiments should be repeatable for performance assessment. A third and most important one is design of experiments that should have controllable parameters to generate desired ground truth to address issues to be investigated. Chapter 4 takes advantage of real image scenes available on web site to simulate synthetic images with various scenarios that can be designed for this purpose.

Chapter 5 is Virtual Dimensionality of Hyperspectral Data that revisits a recently developed concept called virtual dimensionality defined in Chapter 17 of Chang (2003a) as the number of spectrally distinct signatures in hyperspectral imagery. VD has been found to be very useful in many applications (Chang, 2006a, 2006b) such as DR in Wang and Chang (2006a), BS in Chang and Wang (2006), and endmember extraction in Wang and Chang (2006b). Accordingly, a new way of reinterpreting VD becomes imperative. Chapter 5 is a result of such an effort where VD is explored for new interpretation and various techniques are also developed to estimate VD for different applications.

Chapter 6 is Data Dimensionality Reduction. It provides a comprehensive study and survey on many popular and commonly used dimensionality reduction (DR) techniques, which can be treated in two separate categories: dimensionality reduction by transform (DRT) and DR by band selection (DRBS). Specifically, DRT comprises two types of transforms: component analysis (CA)-based transforms, which are derived from statistics of various orders including 2nd order statistics-based principal components analysis (PCA), 3rd order statistics-based skewness, 4th order statistics-based kurtosis and statistical independency-based independent component analysis (ICA) and feature extraction (FE)-based transforms, Fisher's ratio-based linear discriminant analysis (FLDA), and linear mixture model-based OSP. As an alternative to DRT, DRBS selects an appropriate subset of bands from the original band set to replace the high-dimensional original data set with a low-dimensional data set represented by selected bands. So, technically speaking, DRBS performs data reduction, not data compression, by reducing band dimensionality without processing data in the sense that selected bands form a new data cube with all the unselected bands being discarded. While both DRT and DRBS accomplish the same goal, they present different rationales in DR. The former is developed to compact data information in low dimensions via a transform, while the latter represents the original high-dimensional data by its low-dimensional data via band selection. As a consequence, the effectiveness of DR and BS is measured by the transform used for DR and criteria used for BS. Nevertheless, DRT and DRBS do share the same fundamental issue, that is, "how many dimensions are required to be retained after DRT?" and "how many bands are needed for DRBS to faithfully represent the original data?." Interestingly, such an issue has been either overlooked or intentionally avoided in the past because finding an effective criterion for determination of the number of dimensions to be retained or bands to be selected is extremely challenging. Figure 1.1 lists six chapters in Part I to provide background knowledge for follow-up chapters.

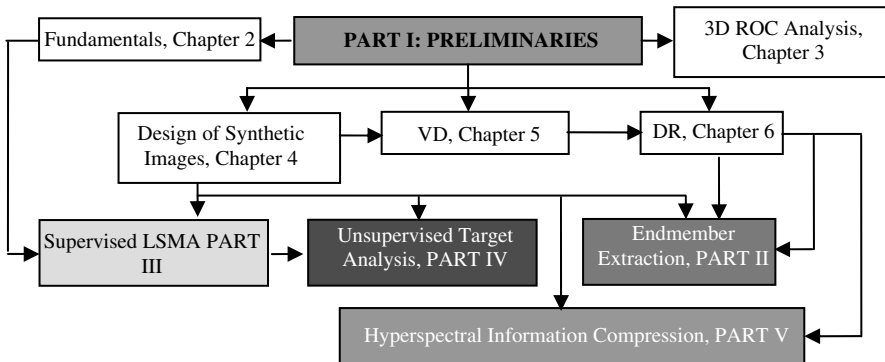


Figure 1.1 Six chapters in Part I to provide background knowledge.

1.5.2 Part II: Endmember Extraction

Endmembers are probably one of most important features in hyperspectral data exploitation since they represent pure signatures used to specify distinct spectral classes. So, finding endmembers becomes a very crucial preprocessing step for hyperspectral image analysis. This is particularly true for linear spectral mixture analysis (LSMA) that requires a set of basic material constituents, referred to as image endmembers to form a linear mixing model to unmix data in terms of abundance fractions of these endmembers. However, the prior knowledge of such image endmembers is usually not available *a priori*. Therefore, endmember extraction comes to play a key role in finding such image endmembers. Unfortunately, the research in endmember extraction has not received much attention in early days until recently. This may be partly due to the fact that many research efforts in remote sensing image processing have been directed to design and development of supervised methods where the necessary prior knowledge is assumed provided *a priori*. In this case, there is no need of finding endmembers. Second, because of low spectral or spatial resolution most image pixels appear in a mixed form rather than as pure pixels. Consequently, the presence of endmembers is considered to be very rare. From a land use/land cover's point of view there may be few endmembers that have little impact on image classification. However, from a viewpoint of intelligence endmembers provide crucial and critical information since their existence is unexpected. Specifically, when they appear, only a small population will be present and cannot be identified by prior knowledge. Additionally, the low probability of their occurrence also makes their detection very difficult. Part II is devoted to this topic. Most importantly, it develops various algorithms of different forms for endmember extraction.

Basically, an endmember extraction algorithm (EEA) can be categorized into simultaneous EEA (SM-EEA) and sequential EEA (SQ-EEA) depending upon how it generates endmembers. An SM-EEA generates a required number of endmembers all together compared to an SQ-EEA, which generates one endmember at a time until it reaches a required number of endmembers. On the other hand, based on how initial conditions are used for initialization, an EEA can be also categorized into initialization-driven EEA (ID-EEA) and random EEA (REEA). These two types of EEAs adopt completely opposite philosophies. An ID-EEA selects a specific set of initial endmembers to avoid randomness caused by the use of random initial endmembers compared to an REEA, which converts the disadvantage resulting from random nature of initial endmembers to an advantage of making an EEA immune to random initial conditions. In order to treat EEAs systematically and logically, Chapter 7 first considers SM-EEAs followed by SQ-EEAs in Chapter 8, ID-EEAs in Chapter 9, and REEA in Chapter 10. Finally, Part II is concluded by Chapter 11, which explores relationships among various EEAs studied in Chapters 7–10. Figure 1.2 outlines the organization of five chapters in Part II.

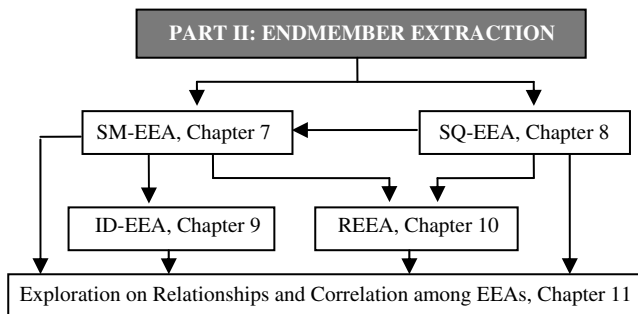


Figure 1.2 Organization of five chapters in Part II.

1.5.3 Part III: Supervised Linear Hyperspectral Mixture Analysis

Supervised linear hyperspectral mixture analysis (SLSMA) is probably the most widely used hyperspectral imaging technique to perform various tasks for data analysis. It makes an assumption that a data sample vector can be described by a linear mixing model as a linear mixture of a finite number of known basic signature constituents called image endmembers, from which it can be unmixed via a specific linear spectral unmixing technique into abundance fractions of these image endmembers. Since SLSMA has been previously treated in the book by Chang (2003a), the five chapters, Chapters 12–15 presented in this book, can be considered as an expansion of SLSMA and complement to the LSMA discussed in Chang (2003a). Chapter 12 revisits the orthogonal subspace projection originally developed by Harsanyi and Chang (1994). In particular, when only partial knowledge such as desired target information is provided with no prior background knowledge, OSP can be implemented as the constrained energy minimization developed in Harsanyi's dissertation (1993). If no prior knowledge is available, then OSP can be implemented as RX detector (Reed and Yu, 1990) for anomaly detection. Chapter 13 presents a third approach to SLSMA, Fisher's linear spectral mixture analysis (FLSMA), which replaces the signal-to-noise ratio criterion used by OSP or least squares error (LSE) used by LSOSP with the criterion of Fisher's ratio. Chapter 14 further extends OSP and FLSMA to WAC-LSMA by replacing the commonly used LSE with weighted LSE. While Chapters 13 and 14 extend SLSMA via imposing constraints on the used linear mixing model, Chapter 15 derives kernel-based LSMA, which extends SLSMA techniques to their kernel-based counterparts via nonlinear functions. Figure 1.3 outlines the organization of four chapters in Part III.

1.5.4 Part IV: Unsupervised Hyperspectral Analysis

One of major tasks in hyperspectral imaging is target detection and classification. Due to its high spectral resolution, targets of interest are generally different from those in multispectral imagery. For example, endmembers and anomalies that generally do not contribute much to land cover/land use classification are actually crucial in hyperspectral image analysis. Other targets of interest in hyperspectral data analysis also include rare minerals in geology, special spices in agriculture and ecology, drug trafficking in law enforcement, combat vehicles in battlefield, man-made targets in intelligent analysis, and so on. Realistically, most of such targets generally appear as either mixed pixels or subpixels. So, the major goal of Part IV is to extend the SLSMA in Part III to ULSMA where two main issues that do not occur in the SLSMA need to be addressed in ULSMA. One is the number of signature sources of interest, p . The other is how to find these signature sources once the value of the p is determined. Since the first issue can be addressed by the concept of VD developed in Chapter 5, the main theme of Part IV is primarily focused on the second issue.

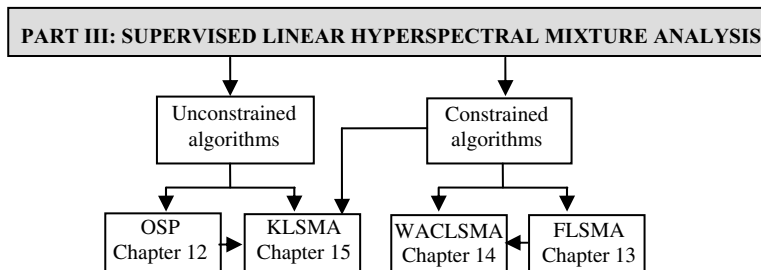


Figure 1.3 Organization of four chapters in Part III.

Chapter 16 investigates two types of hyperspectral measures: signature-based and correlation-weighted measures, both of which can be used to discriminate and identify unknown signature vectors for unsupervised data analysis. The former includes the spectral angle mapper (SAM), Euclidean distance, spectral information divergence (SID), and orthogonal projection divergence (OPD), while the latter uses the sample spectral correlation as a weighting factor to measure signature similarity for discrimination and identification.

Chapter 17 extends SLSMA to ULSMA. In doing so, two approaches are developed to find unknown image endmembers, referred to as virtual signatures (VSs). The first one is to implement LSMA techniques in an unsupervised manner on the original data and its sphered data to find two sets of VSs corresponding to background and target signatures, respectively. A second approach is to use components analysis methods where PCA and ICA are implemented to find unknown background and target signatures, respectively.

Due to substantial amount of information provided by hundreds of contiguous spectral bands it is interesting to know how much information can be extracted from a single hyperspectral image pixel vector as well as how to process the extracted pixel information for data analysis. In traditional image processing the only image pixel information is uniquely specified by its gray-level value. In multispectral image processing with only tens of discrete spectral bands in use, the spectral information provided by a multispectral image pixel is generally very limited compared to that provided by a hyperspectral image pixel. So, the issue in exploration of information extraction from a single hyperspectral image pixel vector has not received much interest as it should. Very little work has been done in the past. For example, an endmember itself provides vital information of a particular spectral class. Another example is an anomaly that provides information in identifying unknown targets. While an endmember is specifically defined, the definition of anomaly seems vague with a general understanding that an anomaly is a target whose spectral signature is distinct from those of pixels in its surrounding neighborhood. However, how large should a surrounding neighborhood be for a pixel vector to be qualified as anomalous pixel vector? So far, there is no answer to it. More generally, for a given pixel vector, how can we characterize the pixel vector as a subpixel vector or a mixed pixel vector or an anomalous pixel vector or a pixel vector of some other type? Besides, can an endmember be a pure pixel vector, in which case it is referred to as endmember pixel vector or vice versa? Can a pixel be both an anomalous pixel vector and an endmember pixel vector? As a complete opposite to anomaly, how can we view a pixel vector if the spectral signatures of pixel vectors in its proximity are very similar and close to each other? Interestingly, these issues have never been investigated on a single pixel vector basis. So, Chapter 18 investigates the issue of “what spectral information can be extracted from a single hyperspectral image pixel vector?” Figure 1.4 outlines the organization of two chapters in Part IV.

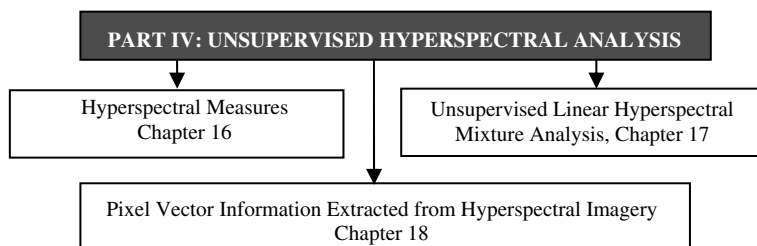


Figure 1.4 Organization of three chapters in Part IV.

1.5.5 Part V: Hyperspectral Information Compression

Data compression has received increasing interest in hyperspectral data analysis because of the vast amount of data volumes needed to be processed and significant redundancy resulting from high interband spectral correlation. Since a hyperspectral image can be viewed as a 3D image cube, a common practice is a direct application of 3D compression techniques available in image/video processing to hyperspectral imagery so as to achieve so-called hyperspectral data compression. Unfortunately, there are several issues arising from such an approach. One is how to deal with spectral compression from very high spectral resolution provided by a hyperspectral imaging sensor. The reason why the hyperspectral imagery is called “*hyperspectral*” is due to its wealthy spectral information, which offers unique spectral characterization that cannot be provided by spatial information, particularly, the spectral profile information provided by sub-pixels and mixed pixels across its acquired wavelength range by hundreds of spectral channels. Therefore, from a hyperspectral imagery point of view, spectral information is usually more important and crucial than spatial information when it comes to hyperspectral image analysts. When hyperspectral compression is performed, extra care must be taken in order to preserve spectral characteristics and properties. For example, when targets of interest are rare such as anomalies and endmembers, their spatial extent is generally very small and limited. Thus, the spatial correlation resulting from such targets will be too little to be used for spatial compression. In this case a direct spatial compression without taking into account spectral properties of these targets may result in significant loss of information that characterizes these targets. As a consequence, blindly applying 3D compression techniques to hyperspectral data may not be able to achieve effective compression from an exploitation perspective. Accordingly, a more appropriate approach is to consider “*information*” compression rather than “*data*” compression since the compression is performed based on preservation of the information of interest instead of reduction in data size. More specifically, an effective technique in compressing data size does not necessarily imply that it is also effective in compressing information to be retained. To resolve this dilemma, an effective means of compressing hyperspectral imagery may be one that performs compression in a two-stage process that carries out spectral compression in the first stage to preserve crucial spectral information to avoid being compromised by the follow-up spatial compression in the second stage (Ramakrishna *et al.*, 2005a, 2005b). Such a two-stage compression is referred to as *hyperspectral information compression* or *exploitation-based lossy hyperspectral data compression* in this book as opposed to *lossy hyperspectral data compression*, commonly referred in the literature. Five chapters are presented in Part V and outlined in Figure 1.5.

Chapter 19 reviews issues arising in data compression commonly used in the literature and further introduces a new concept of hyperspectral information compression or exploitation-based lossy hyperspectral data compression where various approaches can be derived for different applications in data exploitation. This chapter is followed by two new approaches to hyperspectral information compression developed in Chapters 20 and 21, which develop techniques to process spectral dimensions and band dimensions in a progressive manner, referred to as progressive spectral dimensionality process (PSDP) and progressive band dimensionality process (PBDP), respectively. In order to more effectively determine spectral and band dimensionality to be used for material classification Chapter 22 presents a new idea of dynamic dimensionality allocation (DDA). By taking advantage of PBDP in Chapter 21 and DDA in Chapter 22 a new approach to band selection, called progressive band selection (PBS), is further developed and presented in Chapter 23.

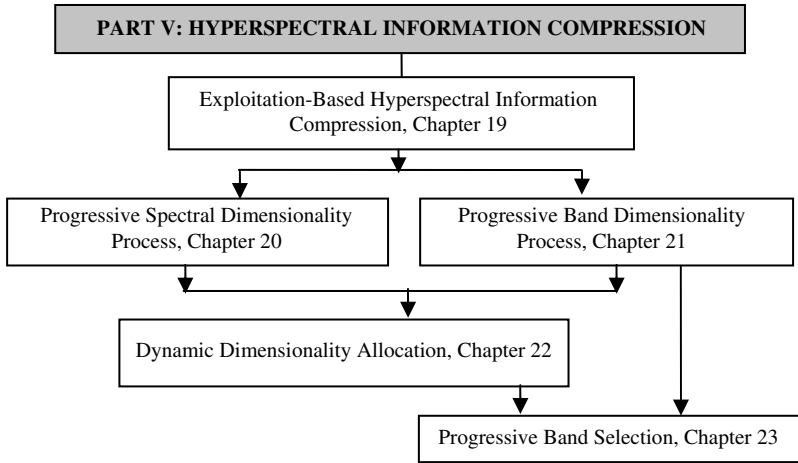


Figure 1.5 Organization of three chapters in Part V.

1.5.6 Part VI: Hyperspectral Signal Coding

So far, data processing discussed in all the previous chapters, Chapters 7–23, is considered as hyperspectral image processing because the considered data are image data cubes and the techniques are developed to process hyperspectral data as an image cube with data samples treated as image pixel vectors. However, due to the use of hundreds of spectral channels a hyperspectral data sample vector already contains spectral information that can be used for data analysis without relying on sample spectral correlation provided by image structures. So, instead of considering a data sample vector as an image pixel vector in an image cube, a data sample vector can also be processed as a one-dimensional signal, referred to as a hyperspectral signal or signature vector rather than as a hyperspectral image pixel vector. In this case a hyperspectral signal is a spectral signature of a material substance specified by hundreds of spectral channels across a certain range of wavelengths. In this book, both hyperspectral signal and signature vector will be used interchangeably as appropriate. The data processing of hyperspectral signals or signature vectors is called hyperspectral signal processing to distinguish it from hyperspectral image processing discussed in previous chapters. The only difference between hyperspectral image processing and hyperspectral signal processing is that the former takes advantage of statistics resulting from spectral correlation among pixel vectors in an image cube, while the latter processes a hyperspectral signal as an individual 1D signal such as signatures from spectral libraries or databases without accounting for spectral correlation among sample signals. As a result, when a hyperspectral signal is processed, the information available for processing is only the spectral information within the signal without referencing spectral correlation with other signals. Accordingly, 1D hyperspectral signal processing is primarily used as signal discrimination, detection, classification, representation, and identification. Having this clear distinction in mind, Part VI and Part VII are devoted to hyperspectral signal processing with an understanding that no sample spectral correlation is available to be used for data processing.

The main focus of Part VI is on signal coding that encodes a hyperspectral signal as a code word for its discrete representation. How fine and accurate such discrete representation of a hyperspectral signal can be is determined by the total number of bits used for encoding. Three types of encoding methods are developed in this part. One is binary coding in Chapter 24, which performs

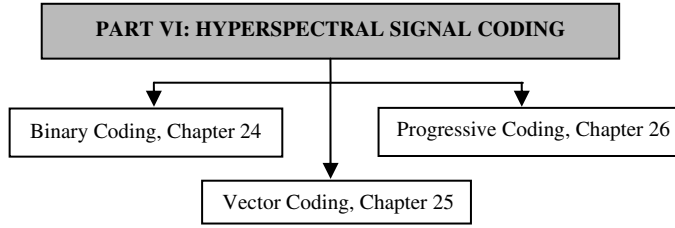


Figure 1.6 Organization of three chapters in Part VI.

memoryless coding. Another is vector coding in Chapter 25, which takes advantage of memory to perform signature coding. A third one discussed in Chapter 26 is progressive coding, which encodes a hyperspectral signal stage by stage in a progressive manner. Figure 1.6 outlines the organization of three chapters in Part VI.

1.5.7 Part VII: Hyperspectral Signal Feature Characterization

While the hyperspectral signal coding considered in Part VI converts a hyperspectral signal to a codeword as its discrete representation so that different hyperspectral signatures can be discriminated and identified via their encoded code words, Part VII can be considered as a counterpart of Part VI to perform hyperspectral signal characterization by converting a hyperspectral signal as a continuous representation. Three major techniques are developed: OSP-based variable-number variable-band selection (VNVBS) in Chapter 27 for hyperspectral signals, Kalman filter-based techniques in Chapter 28 for hyperspectral signal estimation, and wavelet-based techniques in Chapter 29 for hyperspectral signal representation. Figure 1.7 outlines the organization of these three chapters in Part VII.

1.5.8 Applications

This book concludes with applications of hyperspectral data processing in various areas.

1.5.8.1 Chapter 30: Applications of Target Detection

The subpixel target detection discussed in Chapter 2 has major interests in many applications. Since the size of a subpixel target is smaller than pixel resolution specified by ground sampling distance, it is embedded in a single pixel vector and cannot be visualized by inspection. Therefore,

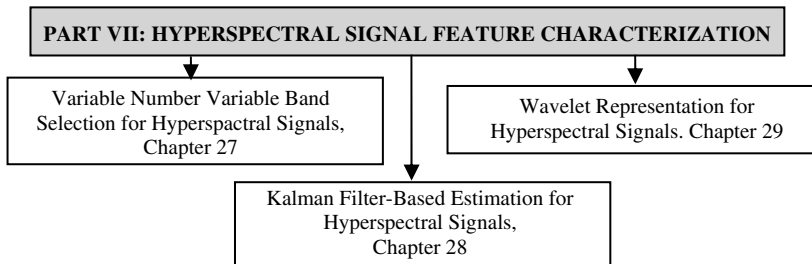


Figure 1.7 Organization of three chapters in Part VII.

it looks like that the best we can do for a subpixel target is detection and finding the size of a subpixel target seems out of reach. Chapter 30 provides a means of doing so. Specifically, the size of a subpixel target can be calculated by multiplying the pixel resolution with the estimated abundance fraction of the subpixel target embedded in a pixel vector. Consequently, finding the true size of a subpixel target is equivalent to accurately estimating the abundance fraction of a subpixel target.

Many problems addressed by target detection assume that the targets to be detected are exposed, in which case it makes detection easy and more effective. However, in remote sensing targets of interest may be hidden under natural environments due to terrain characteristics such as shadow and shade. On the other hand, in many military and intelligence applications, the targets of interest may be concealed weapons or combat vehicles, which are camouflaged or canvassed. Detecting such concealed targets generally presents a great challenge in an unknown image scene due to the fact that the prior knowledge about targets of interest and background is not available. The second part of Chapter 30 develops an approach to detection of unknown concealed targets. It comprises three successive stage processes: (1) band selection procedure in the first stage; (2) band ratio approach in the second stage; and (3) automatic target detection in the third stage. The objective of the band selection is to select an appropriate set of band images for the band ratio transformation and the selected bands are subsequently ratioed to form a desired set of images used for subsequent automatic target detection carried out in the third stage.

1.5.8.2 Chapter 31: Nonlinear Dimensionality Expansion to Multispectral Imagery

The data processing techniques developed in this book are mainly derived from a perspective of how to process hyperspectral imagery. Their applications to multispectral imagery may not be immediately obvious and trivial. Specifically, the pigeon-hole principle described in Section 1.3 that holds for hyperspectral imagery is no longer true for multispectral imagery and virtual dimensionality. In order for a hyperspectral imaging technique to be applied to multispectral imagery, it hinges on two key issues, how to define a hyperspectral image and a multispectral image as well as how to distinguish one from another. Interestingly, the pigeon-hole principle once again proves to be a valuable means of doing so. When there are few pigeon holes than pigeons, it implies that few spectral bands than signal sources can be used for signal discrimination in which case the image is defined as a multispectral image. Otherwise, it is a hyperspectral image. Such definitions seem controversial in the first place. As a matter of fact, similar definitions can be found in ICA. That is, if the number of data sample vectors is fewer than the number of signal sources to be separated, an ICA is defined as an over-complete ICA. Otherwise, an ICA is defined as an under-complete ICA. The definitions of over-complete ICA and under-complete ICA shed light on how to distinguish multispectral image from hyperspectral images. In ICA a data sample vector represents a linear mixture of random signal sources to be separated. This is similar to viewing a data sample vector as a linear mixture of signal sources to be present in the data. So, LSMA used to unmix a multispectral image tries to solve an over-complete linear spectral unmixing problem, while LSMA used to unmix a hyperspectral image intends to solve an under-complete linear spectral unmixing problem. By virtue of this interpretation, this chapter develops two approaches to conversion of a hyperspectral imaging technique to a multispectral imaging technique by nonlinear dimensionality expansion (NDE). One is band dimensionality expansion, which implements a band expansion process (BEP) to create new additional images from the original set of spectral images via nonlinear functions. The other is kernel-based method that kernelizes LSMA-based techniques via nonlinear kernels to solve linear nonseparability issue arising in multispectral image analysis.

1.5.8.3 Chapter 32: Multispectral Magnetic Resonance Imaging

Recently, a new application of hyperspectral imaging techniques in multispectral imagery, magnetic resonance (MR) image analysis, has been investigated where MR images can be considered as multispectral images and each image acquired by a particular MR pulse sequence can be considered as a spectral band image. As a result, MR images are actually an image cube collected by particularly designed MR image pulse sequences. With this interpretation Chapter 32 extends results in Chapter 31 to MR image analysis.

1.6 Laboratory Data to be Used in This Book

Three sets of laboratory data will be used for experiments in this book, two of which were collected by the airborne visible infrared imaging spectrometer (AVIRIS) and the third one is a gas data set.

1.6.1 Laboratory Data

One data set to be used in this book is the one used in Harsanyi and Chang (1994). It is AVIRIS reflectance data shown in Figure 1.8, which has five field reflectance spectra, blackbrush, creosote leaves, dry grass, red soil, and sagebrush with spectral coverage from 0.4 to 2.5 μm and 158 bands after the water bands are removed.

1.6.2 Cuprite Data

Another useful laboratory data that is available on the web site <http://speclab.cr.usgs.gov/> is the reflectance spectra of five USGS ground-truth mineral spectra: alunite (A), buddingtonite (B), calcite (C), kaolinite (K), and muscovite (M) shown in Figure 1.9. Each of the five mineral spectral signatures is collected by 224 spectral bands at spectral resolution of 10 nm in the range of 0.4–2.5 μm .

1.6.3 NIST/EPA Gas-Phase Infrared Database

A third data set is one provided by the National Institute of Standards and Technology (NIST) and also available on the web sites <http://www.nist.gov/srd/nist35.htm> and webbook.nist.gov/chemistry. This data set was also used for the study in Kwan *et al.* (2006). It contains the nine

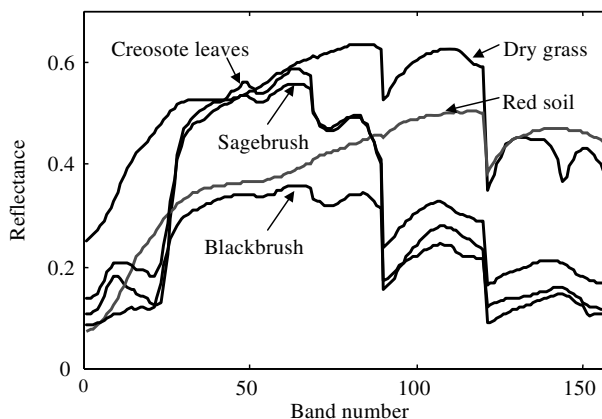


Figure 1.8 Spectra of five AVIRIS reflectances. (See the color version of this figure in Color Plates section).

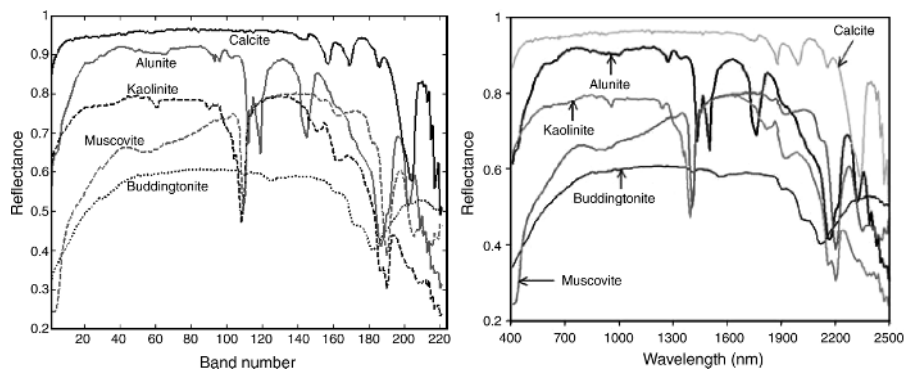


Figure 1.9 Five USGS ground-truth mineral spectra.

Table 1.1 Nine gas agent data signatures

Signature no.	Signature name
s_1	2-Chloroethymethyl sulfide
s_2	Diethyl ethylphosphonate
s_3	Ethanol
s_4	Freon 114
s_5	<i>n</i> -Butyl bromide
s_6	Bis-2-ethyl-1-hexyl phosphonate
s_7	Benzyl benzoate
s_8	Dibenzyl ether
s_9	Piperidine

gas agents labeled by $\{s_i\}_{i=1}^9$ listed in Table 1.1 with their spectral signatures shown in Figure 1.10. This data set is included particularly for signal processing algorithm design and development for hyperspectral signal processing to investigate hyperspectral signature analysis and characterization in Part VI and Part VII.

Except that the frequency range of s_1 is $550\text{--}3846\text{ cm}^{-1}$ acquired by 825 bands, all the $\{s_i\}_{i=2}^9$ has frequency range of $450\text{--}3966\text{ cm}^{-1}$ acquired by 880 bands, giving each signature a spectral resolution of about 4 cm^{-1} per band.

1.7 Real Hyperspectral Images to be Used in this Book

Three real hyperspectral image data sets are frequently used in this book for experiments. Two are AVIRIS real image data sets, Cuprite in Nevada and Purdue’s Indian Pine test site in Indiana. A third image data set is HYperspectral Digital Imagery Collection Experiment (HYDICE) image scene. Each of these three data sets is briefly described as follows.

1.7.1 AVIRIS Data

Two AVIRIS data sets presented in this section are Cuprite data and Purdue’s data, which can be used for different purposes in applications. The Cuprite data set is generally used for endmember

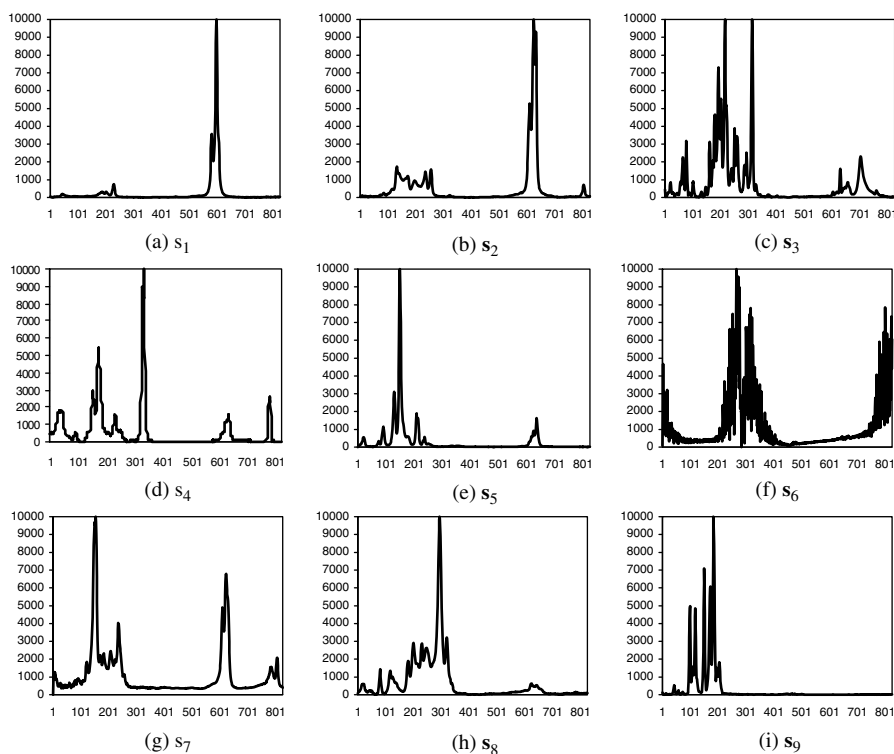


Figure 1.10 Spectral signatures of nine chemical/infrared data signatures. (See the color version of this figure in Color Plates section).

extraction and target detection, while the Purdue's data set is mainly used for land cover/land use classification.

1.7.1.1 Cuprite Data

One of the most widely used hyperspectral image scenes available in the public domain is Cuprite mining site, Nevada, as shown in Figure 1.11(a). It is an image scene of 20 m spatial resolution collected by 224 bands using 10 nm spectral resolution in the range of 0.4–2.5 μm . The center region shown in Figure 1.11(b), cropped from the image scene in Figure 1.10(a), has size of 350×350 pixel vectors.

Since it is well understood mineralogically and has reliable ground truth, this scene has been studied extensively. Two data sets for this scene, reflectance and radiance data, are also available for study. There are five pure pixels in Figure 1.11(a, b) that can be identified to be corresponding to five different minerals, alunite (A), buddingtonite (B), calcite (C), kaolinite (K), and muscovite (M) labeled by A, B, C, K, and M, respectively, in Figure 1.12(b) with their corresponding reflectance and radiance spectra shown in Figure 1.12(c, d).

These five pure pixels are carefully verified using laboratory spectra provided by the USGS (available from <http://speclab.cr.usgs.gov>) and selected by comparing their reflectance spectra in Figure 1.12(c) against the lab reflectance data in Figure 1.9. Figure 1.12(e) also shows an alteration map for some of the minerals, which is generalized from ground map provided by the USGS

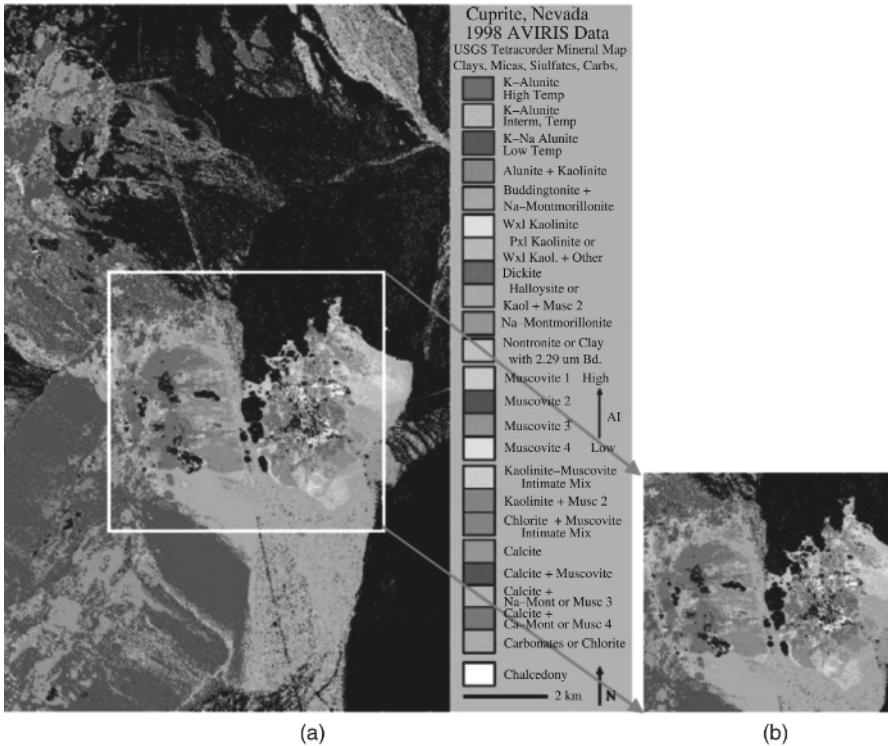


Figure 1.11 Cuprite image scene, (a) original Cuprite image scene; (b) the image cropped from the center region of the original scene in (a) (350×350). (See the color version of this figure in Color Plates section).

and obtained by Tricorder SW version 3.3. It should be noted that this radiometrically calibrated and atmospherically corrected data set available from <http://aviris.jpl.nasa.gov> is provided in reflectance units with 224 spectral channels where the data has been calibrated and atmospherically rectified using the ACORN software package. It is recommended that bands 1–3, 105–115, and 150–170 be removed prior to data processing due to their low water absorption and low SNR. As a result, a total of 189 bands are used for experiments as shown in Figure 1.11(c, d). The steps to produce spectra in Figure 1.12(c, d) can be described as follows:

1. Download from <http://speclab.cr.usgs.gov/> the laboratory reflectance data.
2. Use spectral angle mapper (SAM) as a spectral similarity measure to identify the five pixels in Figure 1.12(a) that correspond to the five reflectances obtained in step 1 by the following procedure:
 - Remove noisy bands from the five reflectance data.
 - Remove bands with abnormal readings from the spectral library.
 - In order to measure spectral similarity, there is still a need of removing several bands to account for compatibility.

It should be noted that the ground truth is not stored in a “file.” The locations of the five minerals are identified by comparing their reflectance spectra against their corresponding lab reflectances in the spectral library.

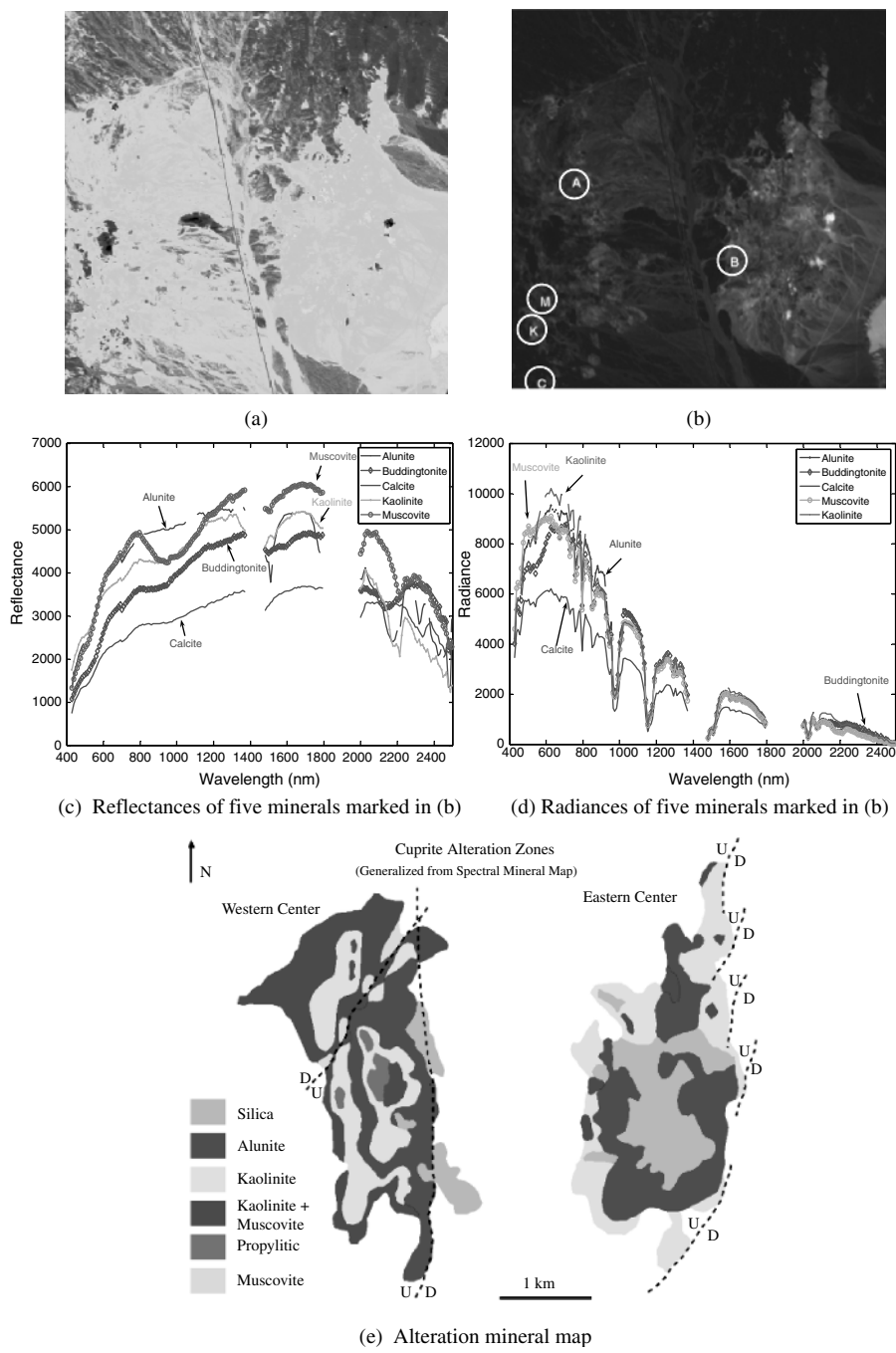


Figure 1.12 (a) Spectral band number 170 of the Cuprite AVIRIS image scene; (b) spatial positions of five pure pixels corresponding to minerals: alunite (A), buddingtonite (B), calcite (C), kaolinite (K), and muscovite (M); (c) reflectances of five minerals marked in (b) in wavelengths; (d) radiances of five minerals marked in (b) in bands; and (e) alteration mineral map available from USGS. (See the color version of this figure in Color Plates section).

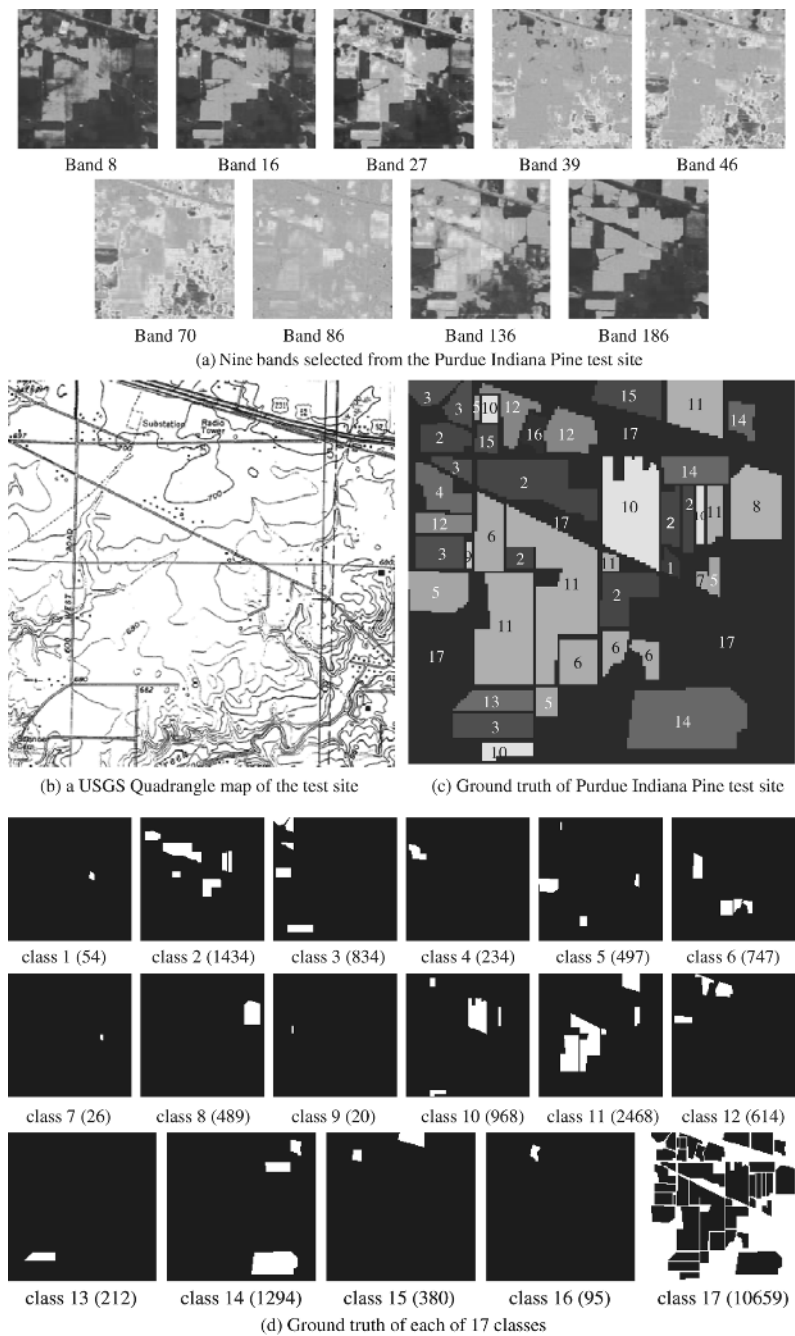


Figure 1.13 AVIRIS image scene: Purdue Indiana Pine test site. (a) Nine bands selected from the Purdue Indiana Pine test site; (b) a USGS quadrangle map of the test site; (c) ground truth of Purdue Indiana Pine test site; and (d) ground truth of each of 17 classes. (See the color version of this figure in Color Plates section).

1.7.1.2 Purdue's Indiana Indian Pine Test Site

Another most widely used real AVIRIS image data set is Purdue's Indiana Indian Pine test site, which has 20 m spatial resolution and 10 nm spectral resolution in the range of 0.4–2.5 μm with size of 145×145 pixel vectors taken from an area of mixed agriculture and forestry in North-western Indiana, USA. The data set is available on the web site <http://cobweb.ecn.purdue.edu/~biehl/MultiSpec/documentation.html> (both download link and ground truth are provided) and was recorded in June 1992 with 220 bands with water absorption bands, bands 104–108 and 150–162 removed and leaving only 202 bands. Figure 1.13(a) shows nine bands selected from the web site and a USGS quadrangle map of the test site provided in Figure 1.13(b).

According to the ground truth provided in Figure 1.13(c) there are 17 classes in this image scene shown in Figure 1.13(d) including the background labeled by class 17, which has a wide variety of targets such as highways, railroad, houses/buildings, and vegetation that may not be of interest in agricultural applications but may be of great interest in other applications such as anomaly detection. The total number of data samples in the scene is $145 \times 145 = 21,025$. Table 1.2 lists labels of each of 17 classes where the numeral in parenthesis under each of 17 classes in Figure 1.13(d) is the number of data samples in that particular class.

Due to the early season of harvest when the data were collected, some cultivated land has very little canopy cover. For example, the corn area can be divided into three classes based on how much is left on the land, which are corn-no till, -min, and corn (class 2–4). The soybean area also can be divided into soybean-no till, -min, and -clean (class 10–12). The grass is mixed with four other materials, which are classified as grass/pasture, grass/trees, grass/pasture-mowed, and bldg-grass-green-drives (class 5, 6, 7, 15). Actually, it is believed that the grass is also mixed in the background. According to Figure 1.13(c, d) (Landgrebe, 2003), the GIS map in Figure 1.13(b) provides the information of “land use” classes instead of “land cover” classes. It means that not every pixel in the map is supposed to be classified into their belonging classes. Additionally, also based on the USGS quadrangle map in Figure 1.13(b), there are dual lane highways (U.S. 52 and 231) and a railroad crossed near the top. The other is Jackson highway, which is near to the middle of the scene. All of them are in the NW–SE direction. Figure 1.13(b) also indicates some houses or buildings by small rectangular dots (Landgrebe, 1998). With this information it is believed that there are at least four classes included in the background: railroad (iron), highway (concrete), houses/buildings (concrete, painted wood, or other materials), and vegetation (grass). The number of classes for such unlabeled areas is important for the unsupervised classification when the total number of classes in the scene is assumed to be unknown.

There are many reasons to select the Purdue Indiana Indian Pine test site for experiments. First of all, it is a well-known image scene available on web site and has been studied extensively. Another is that the pixels in this image scene are heavily mixed. Many algorithms or methods claiming to work well on classification are very likely to break down for this image scene. To the author's best knowledge, most work on this image scene reported in the literature has selected

Table 1.2 Labels of 17 classes

Class 1	Alfalfa	Class 7	Grass/pasture-mowed	Class 13	Wheat
Class 2	Corn-no till	Class 8	Hay-windrowed	Class 14	Woods
Class 3	Corn-min	Class 9	Oats	Class 15	Bldg-grass-green-drives
Class 4	Corn	Class 10	Soybean-no till	Class 16	Stone-steel towers
Class 5	Grass/pasture	Class 11	Soybean-min	Class 17	Background
Class 6	Grass/trees	Class 12	Soybean-clean		

particular areas for study and also supervised based on the provided ground truth. Very little has been done in classification of the entire scene either supervisedly or unsupervisedly. Most interestingly, according to our detailed analysis on the scene, we have found that it is almost impossible to classify all the 17 classes in the image scene even though the complete knowledge of the ground truth provided in Figure 1.13(c, d) is used for classification. This is because pixels in the same class are mixed so badly that values among their spectral signatures measured by any spectral similarity measure vary in a relatively wide range in which pixels in the same class may be classified into different classes and pixels in different classes may be considered to belong to the same class.

From the ground truth provided in Table 1.2, it can be expected that the signatures of three subclasses of corn are close to each other, so are the four subclasses of grass and three subclasses of soybean. However, the relationships among other pairs are still not known. In order to know how much mixing is involved, the signature for each class is calculated by averaging all samples with the same label according to the ground-truth map in Figure 1.12(c). Then the SAM is used to measure how close one class is similar to the other. It has been shown in Liu (2005) that corn and soybean classes (2–4, 10–12) are similar, which account for 6552 pixels, 63% of 10,366 labeled pixels. Similarity also appears in two sets of classes: class 1, 7, 8 and class 6, 9, 13. Surprisingly, the four classes of grass (5, 6, 7, 15), which account for 1650 pixels, are not similar to each other. Additionally, using SAM to measure spectral similarity among 16 classes, it is found that classes 5, 14, 16 seem to be the three most distinct classes and can be classified very easily. It is reasonable and makes sense because class 5 contains chlorophyll, class 14 is wood, and class 16 comprises man-made objects.

With our tremendous experience of working on this image scene, excluding two classes (class 17 that is considered to be the background and class 9 that is considered to be too small) it is found that the spectral signatures of the pixels in the six classes (class 2, class 3, class 4, class 7, class 9, and class 11) are very close in terms of SAM or SID (spectral information divergence in Chang (2003a)). Similarly, the pixels in the three classes (class 8, class 10, and class 15) also have very similar spectral signatures. Hence distinguishing one from another is very difficult. The pixels in the three classes (class 13, class 5, and class 14) have less similar signatures but still present some difficulty with classification. The most dissimilar classes are class 1, class 6, and class 12 that are considered to be easy to classify. By taking into account all the things considered above, we can expect that the classification of this image scene is a great challenge to any hyperspectral imaging algorithm.

1.7.2 HYDICE Data

The HYDICE image scene shown in Figure 1.14(a) has a size of 200×74 pixel vectors along with its ground truth provided in Figure 1.14(b) where the center and boundary pixels of objects are highlighted by red and yellow, respectively. The upper part contains fabric panels with size 3, 2, and 1 m^2 from the first column to the third column. Since the spatial resolution of the data is 1.56 m^2 , the panels in the third column are considered as subpixel objects. The lower part contains different vehicles with sizes of $4 \text{ m} \times 8 \text{ m}$ (the first four vehicles in the first column) and $6 \text{ m} \times 3 \text{ m}$ (the bottom vehicle in the first column) and three objects in the second column (the first two have size of 2 pixels and the bottom one has size of 3 pixels, respectively). In this particular scene, there are three types of targets with different sizes, small-size targets (panels of three different sizes, 3, 2, and 1 m^2), and large-size targets (vehicles of two different sizes, $4 \text{ m} \times 8 \text{ m}$ and $6 \text{ m} \times 3 \text{ m}$ and three objects of 2-pixel and 3-pixel sizes) that can be used to validate and test anomaly detection performance.

Figure 1.14(c) shows an enlarged HYDICE scene from the same flight for visual assessment. It has a size of 33×90 pixel vectors with 10 nm spectral resolution and 1.56 m spatial resolution where five vehicles lined up vertically to park along the tree line in the field where the red (R) pixel

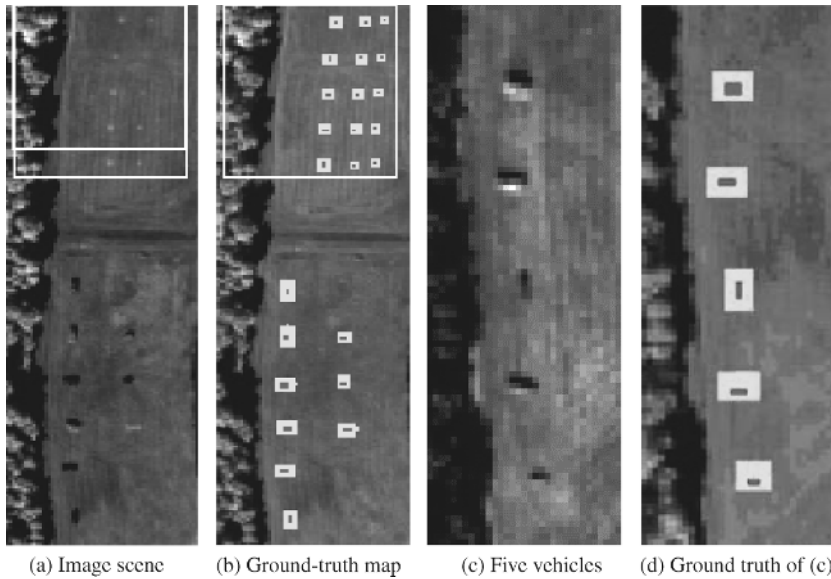


Figure 1.14 HYDICE vehicle scene. (a) Image scene; (b) ground-truth map; (c) five vehicles; and (d) ground truth of (c). (See the color version of this figure in Color Plates section).

vectors (shown as dark pixels) in Figure 1.14(d) show the center pixel of the vehicles, while the yellow (Y) pixels (shown as bright pixels) are vehicle pixels mixed with background pixels.

A third enlarged HYDICE image scene shown in Figure 1.15(a) is also cropped from the upper part of the image scene in Figure 1.14(a, b) marked by a square.

It has a size of 64×64 pixel vectors with 15 panels in the scene. This particular image scene has been well studied in Chang (2003a). Within the scene there is a large grass field background, a forest on the left edge, and a barely visible road running on the right edge of the scene. Low signal/high noise bands: bands 1–3 and bands 202–210; and water vapor absorption bands: bands 101–112 and bands 137–153 were removed. The spatial resolution is 1.56 m, and spectral resolution is 10 nm. There are 15 panels located in the center of the grass field and are arranged in a 5×3 matrix as shown in Figure 1.15(b), which provides the ground-truth map of Figure 1.15(a). Each element in this matrix is a square panel and denoted by p_{ij} with row indexed by $i = 1, \dots, 5$ and column indexed by $j = 1, 2, 3$. For each row $i = 1, \dots, 5$, the three panels were painted by the same material but have three different sizes. For each column $j = 1, 2, 3$, the five panels have the same size but were painted by five different materials. It should be noted that the panels in rows 2 and 3 are made by the same material with different paints, so did the panels in rows 4 and 5. Nevertheless, they were still considered as different materials. The sizes of the panels in the first, second, and third columns are $3 \text{ m} \times 3 \text{ m}$, $2 \text{ m} \times 2 \text{ m}$, and $1 \text{ m} \times 1 \text{ m}$, respectively. So, the 15 panels have 5 different materials and 3 different sizes. Figure 1.15(b) shows the precise spatial locations of these 15 panels where red pixels (R pixels, i.e., dark pixels) are the panel center pixels and the pixels in yellow (Y pixels, i.e., bright pixels) are panel pixels mixed with background. The 1.56 m spatial resolution of the image scene suggests that the panels in the second and third columns, denoted by p_{12} , p_{13} , p_{22} , p_{23} , p_{32} , p_{33} , p_{42} , p_{43} , p_{52} , p_{53} in Figure 1.15(b) are one pixel in size. Additionally, except the panel in the first row and first column, denoted by p_{11} which also has a size of one pixel, all other panels located in the first column are two-pixel panels, which are the panels in the second row with two pixels lined up vertically, denoted by

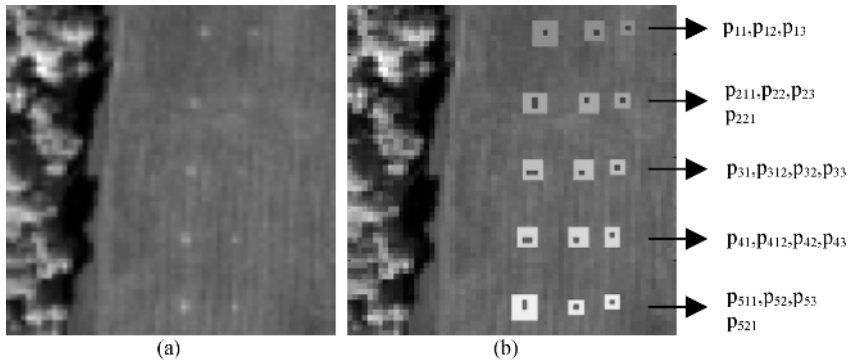


Figure 1.15 (a) A HYDICE panel scene that contains 15 panels; (b) ground-truth map of spatial locations of the 15 panels. (See the color version of this figure in Color Plates section).

p_{211} and p_{221} ; the panel in the third row with two pixels lined up horizontally, denoted by p_{311} and p_{312} ; the panel in the fourth row with two pixels also lined up horizontally, denoted by p_{411} and p_{412} ; and the panel in the fifth row with two pixels lined up vertically, denoted by p_{511} and p_{521} . Since the size of the panels in the third column is $1\text{ m} \times 1\text{ m}$, they cannot be seen visually from Figure 1.15(a) due to its size being smaller than the 1.56 m pixel resolution.

Figure 1.16 plots the five panel spectral signatures obtained from Figure 1.15(b), where the i th panel signature, denoted by p_i , was generated by averaging the red panel center pixels in row i . These panel signatures will be used to represent target knowledge of the panels in each row.

According to visual inspection and ground truth in Figure 1.15(a, b) there are also four background signatures shown in Figure 1.17, which can be identified and marked by interferer, grass, tree, and road. These four signatures along with five panel signatures in Figure 1.16 can be used to form a 9-signature matrix for a linear mixing model to perform supervised linear spectral mixture analysis.

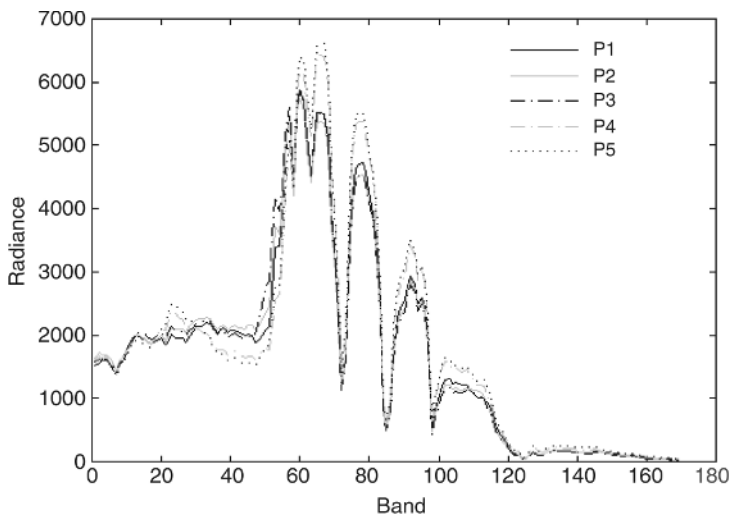


Figure 1.16 Spectra of p_1 , p_2 , p_3 , p_4 , and p_5 .

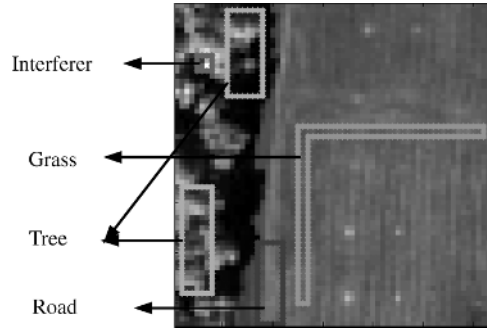


Figure 1.17 Areas identified by ground truth and marked by three background signatures, grass, tree, road plus an interferer. (See the color version of this figure in Color Plates section).

1.8 Notations and Terminologies to be Used in this Book

Since this book primarily deals with real hyperspectral data, the image pixels are generally mixed and not necessarily pure. The term “endmember” is not used here; instead, a general term “signature” or “signature vector” is used. In addition, because we are only interested in target analysis, the term “targets” instead of “materials” is also used throughout this book. In order to make a distinction between a target pixel and its spectral signature vector, we use notation “ $\mathbf{t}$ ” to represent the target pixel vector, “ $\mathbf{r}$ ” for an image pixel vector, and “ $\mathbf{s}$ ” or “ $\mathbf{m}$ ” to indicate its spectral signature vector. We also use bold uppercase for matrices and bold lowercase for vectors. The *italic* upper case “ L ” will be used for the total number of spectral bands, $\mathbf{K}$ for the sample spectral covariance matrix, and $\mathbf{R}$ for the sample spectral correlation matrix. Also, $\delta^*(\mathbf{r})$ is used to represent a detector or classifier that operates on an image pixel vector $\mathbf{r}$ where the superscript “*” in $\delta^*(\mathbf{r})$ specifies what type of a detector or classifier to be used. It should be noted that $\delta^*(\mathbf{r})$ is a real-valued function that takes a form of inner product of a filter vector $\mathbf{w}$ with $\mathbf{r}$, that is, $\delta^*(\mathbf{r}) = (\mathbf{w}^*)^T \mathbf{r}$ with the filter vector $\mathbf{w}^*$ specified by a particular detector or classifier. We also use “ α ” and $\hat{\alpha}$ to represent the abundance vector and its estimate where the notation “hat” over “ α ” indicates “estimate.”

α : Abundance vector

$\hat{\alpha}$: Estimate of the abundance vector α

α_j : j th abundance fraction

$\hat{\alpha}_j$: Estimate of the j th abundance fraction, α_j

$\mathbf{A}$: Weighting or mixing matrix

A_z : Area under an ROC curve

$\mathbf{B}_l$: l th band image

$\mathbf{b}_l$: l th band image represented as a vector

C : Total number of classes

C_j : j th class

$\mathbf{d}$: Desired signature vector

$\mathbf{D}$: Desired signature matrix

$\mathbf{D}_\lambda$: Eigenvalue diagonal matrix

δ : Detector or classifier

Δ : Database

ϵ : Error threshold

$\mathbf{e}_j$: j th endmember

I: Identity matrix
 $I(x)$: Self-information of x
 $k(.,.)$: Kernel function
 K : Total number of skewers used in PPI
K: Sample covariance matrix
 λ : Eigenvalue of sample covariance matrix, **K**
 $\hat{\lambda}$: Eigenvalue of sample correlation matrix, **R**
 l : Index of band number
 L : Total number of spectral channels or bands
 Λ : Eigenvector matrix
 μ : Global sample mean
 μ_j : Global mean of the j th class
 $m(.,.)$: Spectral measure
 $\mathbf{m}_j$: j th signature vector
M: Signature or endmember matrix
n: Noise vector
 N : Total number of image pixel vectors in a band image, i.e., $N = n_r n_c$
 n_c : Number of columns in a band image of a hyperspectral image
 n_D : Number of desired signatures in **D**
 n_r : Number of rows in a band image of a hyperspectral image
 n_{II} : Number of interferers
 n_T : Number of training samples
 n_U : Number of undesired signatures in **U**
 n_{VD} : Value estimated by the VD
 p : Number of endmembers
 P_D : Detection power or probability
 P_F : False alarm probability
 $P_U^\perp$: Projector to reject undesired target signatures in **U**
 q : Number of dimensions to be retained after dimensionality reduction
 $\hat{q}$: Number of spectral bands required to be selected by band selection
 $\mathbf{r}$: Image pixel vector
R: Sample correlation matrix
 σ^2 : Variance
 S_B : Between-class scatter matrix
 S_W : Within-class scatter matrix
 $\mathbf{t}$: Target signature
 τ : Threshold
 $\mathbf{w}$: Weight vector
W: Weight matrix
U: Undesired signature matrix
 $\mathbf{v}$: Eigenvector
 VD_* : The value of the VD obtained by the criterion specified by algorithm “*”
 ξ : Transform used to perform dimensionality reduction
 Ψ : Interference matrix
 $\mathbf{z}$: Projection vector
 $\langle ., . \rangle$: Inner product