Target recognition in hyperspectral images

Amir Z. Averbuch Michael V. Zheludev Valery Zheludev
School of Computer Science
Tel Aviv University, Tel Aviv 69978, Israel

Abstract

We present new algorithms that perform unmixing in hyperspectral images and then recognize targets whose spectral signatures are given. The target can occupy sub- or above pixel. These algorithms combine ideas from algebra and probability theory. Experimental results demonstrate the efficiency and the robustness of these algorithms on real hyperspectral data.

1 Introduction

1.1 Data representation and extraction of spectral information

We assume that an hyperspectral signature of a sought after material is given. In many applications, a fundamental processing task is to automatically identify pixels whose spectra have a specified given spectral shape (signature). This problem raises the following issues: How the measured spectrum of a ground material is related to a given "pure" the spectrum and how to compare between them to determine if they are the same? As a result of spatial and spectral sampling, airborne hyperspectral imaging sensors produce a 3D data structure referred to as a data-cube.

The observed spectral radiance data, or the derived surface reflectance data, can be viewed as a scattering of points in a K-dimensional Euclidean space $\mathbb{R}^K$, where K is the number of spectral bands. Each spectral band is assigned to one axis of the space. All the axes are mutually orthogonal. Therefore, the spectrum of each pixel can be viewed as a vector $x = (x_1, x_2, \ldots, x_k)$ where its Cartesian coordinates x_i are the radiance or the reflectance values at each spectral band (wavelength). Since each component $x_i \geq 0$, then the spectral vectors lie inside a positive cone in $\mathbb{R}^K$. Changes in the level of illumination can change the length of the spectral vector but not its orientation, which is related to the shape of the spectrum. When targets are too small to be resolved spatially or when they are partially obscured or of an unknown shape, then the detection

has to rely on the available spectral information. Unfortunately, a perfect fixed spectrum for any given material does not exist.

Spectra of the same material are probably never identical even in laboratory experiments. This is due to variations in the material surface. The amount of variability is even more profound in remote sensing applications because of the variations in atmospheric conditions, sensor noise, material composition, location, surrounding materials and other factors. As a result, measured spectra, which correspond to pixels with the same surface type, exhibit an inherent spectral variability that prevents the characterization of homogeneous surface materials by unique spectral signatures.

Another significant complication arises from the interplay between the spatial resolution of the sensor and the spatial variability present in the observed ground scene. A sensor integrates the radiance from all the materials within the ground surface that are "seen" by the sensor as a single image pixel. Therefore, depending on the spatial resolution of the sensor and the distribution of surface materials within each ground resolution cell, the result is a hyperspectral data-cube comprised of "pure" and "mixed" pixels, where a pure pixel contains a single surface material and a mixed pixel contains multiple (superposition of) materials. The most widely used spectral mixing model is the linear mixing model, which assumes that the observed reflectance spectrum, for a given pixel, is generated by a linear combination of a small number of unique constituent deterministic spectral signatures known as endmembers. This model is defined with constraints in the following way:

$$x = \sum_{k=1}^{M} a_k s_k + w = Sa + w, \quad \sum_{k=1}^{M} a_k = 1 \text{ additivity constraint}, \quad a_k \geq 0 \text{ positivity constraint} \tag{1.1}$$

where $s_1, s_2, \ldots, s_M$ are the M endmember spectra which are assumed to be linearly independent, $a_1, a_2, \ldots, a_M$, are the corresponding abundances (cover material fractions), and w is an additive-noise vector. Endmembers may be obtained from spectral libraries, in-scene spectra, or geometrical techniques.

1.2 Motivation and research approach

The new methods in this paper are utilized to achieve targets identification with known spectra. Target identification in hyperspectral has the following consecutive steps:

1. Finding suspicious points: there are points whose spectra are different in any norm from the spectra of the points in its neighborhood;

2. Extracting from the suspicious points the spectra of the independent components (unmixing) where one of them is the target that its spectrum fits the given spectrum.

We assume that the spectra of different materials are dependent and unmixing between them depends on the behavior of the first and second derivatives in certain sections. If they are independent, then all the related works such as Maximum Likelihood (ML) and Geometrical approach, which will be mentioned in section 1.3, work well.

The hyperspectral images, which are used in the paper, were captured by the camera AISA Airborne Hyperspectral Systems manufactured by the Specim company. The hyperspectral images were taken from a plane that flew in 10,000 feet height.

1.3 Related work

Linear approach: Under the linear mixing model where the number of endmembers and their spectral signatures are known, hyperspectral unmixing is a linear problem, which can be addressed, for example, by the ML setup [13] and by the constrained least squares approach [7]. These methods do not supply sufficientl accurate estimates and do not reflect the physical behavior. Distinction between different material's spectra is conditioned generally by the distinction in the behavior of the first and the second derivatives and not by a trend.

Independent component analysis (ICA) is an unsupervised source separation process that finds a linear decomposition of the observed data yielding statistically independent components [15, 16]. It has been applied successfully to blind source separation, to feature extraction and to unsupervised recognition .

If the mixture of hyperspectral data are linear, then ICA is a possible tool for unmixing. For example, the application of ICA to hyperspectral data was proposed in [17], where the endmember signatures are treated as sources and the mixing matrix is composed by the abundance fractions - see [18, 16, 27, 19, 20, 21], where the sources are the abundance fractions of each endmember. The first approach has two difficulties: 1. The number of samples are limited to the number of channels. 2. The process of pixel selection, which plays the role of mixed sources, is not straightforward. The second approach also faces difficulties since the sum of the abundance fractions is constant. This implies statistical dependency among abundances (i.e. among sources). This dependency violates a key assumption of ICA of having statistical source independency. The applicability of ICA to hyperspectral images is thus compromised. In addition, hyperspectral data are immersed in noise, which degrades the ICA performance. Independent factor analysis (IFA) [22, 23] was introduced as a method for recovering independent hidden sources from their observed noisy mixtures. IFA implements two steps: 1. Source densities and noise covariance are estimated from the observed data by ML; 2. Sources are reconstructed by an optimal nonlinear estimator. Although IFA is a well

suited technique to unmix independent sources under noisy observations, the dependency among abundance fractions in hyperspectral imagery compromises, as in the ICA case, the IFA performance.

The impact of source dependency on unmixing hyperspectral data with ICA and IFA algorithms is investigated in [24]. It shows that these algorithms do not correctly unmix hyperspectral data such that the unmixing matrix, which minimizes the mutual information, can be very far from being true.

Geometric approach: Assuming a linear mixing scenario where each observed spectral vector is given by

$$\mathbf{r} = \mathbf{x}+\mathbf{n}=\mathbf{M}\gamma\mathbf{a}+\mathbf{n}, \quad \gamma\mathbf{a} = \mathbf{s}, \tag{1.2}$$

where $\mathbf{r}$ is an L-vector (L is the number of bands), $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \dots \mathbf{m}_p]$ is the mixing matrix ($\mathbf{m}_i$ denotes the ith endmember signature and p is the number of endmembers present in the sensed area), $\mathbf{s} \stackrel{\Delta}{=} \gamma\mathbf{a}$ (γ is a scale factor that models illumination variability due to a surface topography), $\mathbf{a} = [a_1, a_2, \dots a_p]^T$ is the abundance vector that contains the fractions of each endmember ($(\cdot)^T$ denotes a transposed vector) and $\mathbf{n}$ models is the system's additive noise. Owing to physical constraints, abundance fractions are non-negative and satisfy the so-called positivity constraint $\sum_{k=1}^{p} a_k = 1$. Each pixel can be viewed as a vector in a L-dimensional Euclidean space, where each channel is assigned to one axis. Since the set $\{\mathbf{a} \in \mathbb{R}^p : \quad \sum_{k=1}^{p} a_k = 1,\ a_k > 0 \text{ for all } k\}$, is a simplex, then the set $S_x \stackrel{\Delta}{=} \left\{\mathbf{x} \in \mathbb{R}^L : \quad \mathbf{x} = \mathbf{M}\mathbf{a}, \quad \sum_{k=1}^{p} a_k = 1,\ a_k > 0 \text{ for all } k\right\}$, is also a simplex whose vertices correspond to endmembers.

Several approaches [1, 2, 3] exploited this geometric feature of hyperspectral mixtures. The minimum volume transform (MVT) algorithm [3] determines the simplex of a minimal volume that contains the data. The method presented in [4] is also of MVT type, but by introducing the notion of bundles, it takes into account the endmember variability that is usually present in hyperspectral mixtures.

The MVT type approaches are complex from the computational point of view. Usually, these algorithms first find the convex hull defined by the observed data and then fit a minimum volume simplex to it. Aiming at a lower computational complexity, some algorithms such as the pixel purity index (PPI) [2] and the N-FINDR [5] still find the minimum volume simplex that contains the data cloud. They assume the presence in the data of at least one pure pixel of each endmember. This is a strong assumption that may not be true in many datasets. In any case, these algorithms find the set of most pure pixels in the data.

Extending subspace approach: A fast unmixing algorithm, termed *vertex component analysis* (VCA) is described in [6]. The algorithm is unsupervised and exploits two facts: 1. The endmembers are the vertices of a simplex; 2. The affine transformation of a simplex is also a simplex. It works with projected and with unprojected data. As PPI and N-FINDR algorithms, VCA also assumes the presence of pure pixels in the data. The algorithm iteratively projects data onto a direction orthogonal to the subspace spanned by the endmembers already determined. The new endmember's signature corresponds to the extreme of the projection. The algorithm iterates until all the endmembers are exhausted. VCA performs much better than PPI and better than or comparable to N-FINDR. Yet, its computational complexity is between one and two orders of magnitude lower than N-FINDR.

If the image is of size approximately 300×2000 pixels, then this method, which builds linear span in each step, is too computational expansive. In addition, it relies on "pure" spectra which are not available all the time.

The paper has the following structure: Section 2 presents an algorithm that identifies the target's spectrum when the target occupies whole pixel/s without mixing with the background. This method is needed when the target's spectrum is distorted by atmosphere conditions and noised. This method performs better than the application of a correlation comparison. Section 3 presents an unmixing method that is based on neighborhood analysis of each pixel. This algorithm contains two parts. In the first part, the suspicious points are discovered. The algorithm is based on the properties of connected components of pixels from the neighborhood which are correlated with the current pixel. The second part unmixes a suspicious point. It is based on a projection to the orthogonal complement of the linear span of the neighboring background spectra. An alternative unmixing algorithm is given in section 4. This algorithm does not necessitate an expensive computational procedure for neighborhood analysis. It requires that the correlation of the second derivative of the spectra from different materials be close to zero.

2 Method I: Weak dependency recognition (WDR) of targets that occupy one or more pixels

When a target occupies one or more pixels, the target's spectrum can be recognized by comparing it with with patterns from a given database of spectra. An initial optional step is to denoise the image by the application of a moving average method. Then, we determine if the given target's spectrum and the spectrum of the current pixel are dependent. This procedure is described next.

Two discrete functions Y_1 and Y_2 are weakly dependent if there is a monotonous function F such that $Y_1 = F(Y_2)$.

Let T be a given target's spectrum and P is the current pixel's spectrum. We consider the spectra of T and P as discrete vectors. In general, we assume that T and P are normalized and centralized. The following hypotheses are assumed:

H_0: T and P are weakly dependent.

H_1: T and P are not weakly dependent.

2.1 Hypotheses check

We find the orthogonal transformation that permutes the coordinates of T into a decreasing order. This permutation Π is applied to P and T. We get that $P_1 = \Pi(P)$, $T_1 = \Pi(T)$ where T_1 is monotonic. If H_0 holds, which means that T and P are weakly dependent, then the values of P_1 are either monotonic decreasing or increasing and the first and second derivatives of P_1 are close to zero - see Fig. 2.1 (top and bottom left, respectively). Otherwise, H_1 holds and P_1 has an oscillatory behavior - see Fig. 2.1 (top right). In addition, P_1 has a subset of coordinates whose first and second derivatives have an oscillatory behavior - see Fig. 2.1 (bottom right).

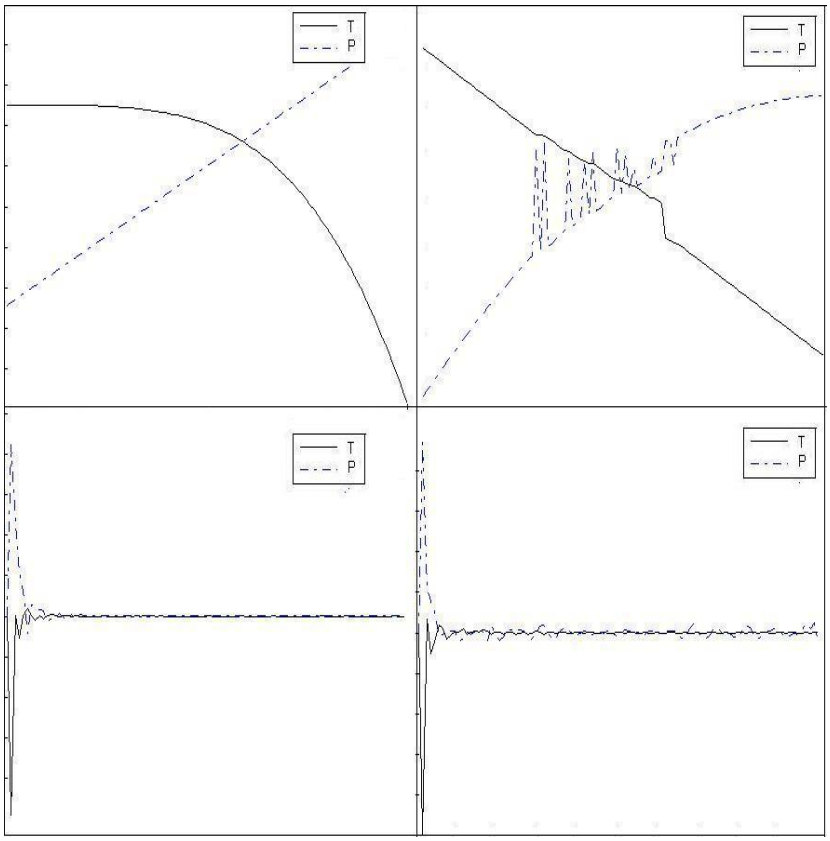

Figure 2.1: Left: Weak dependency between T and P. Right: Not weak dependency

If the permutation of the coordinates of P provides monotonic (decrease or increase) behavior for the values of P, then the first and second derivatives of P have a minimal norm. This is another criterion for having weak dependency.

Denote the second derivative of P by P_2. Let $n(P) = \|P_2\|_\infty$. Let $D = 0.2$ be a threshold that was determined experimentally for normalized and centralized vectors. If $n(P) > D$ then H_1 holds. If $n(P) \leq D$ then H_0 holds. Only weakly dependent pixels contain targets.

2.2 Experimental results

Figures 2.2-2.4 display the results after the application of this algorithm.

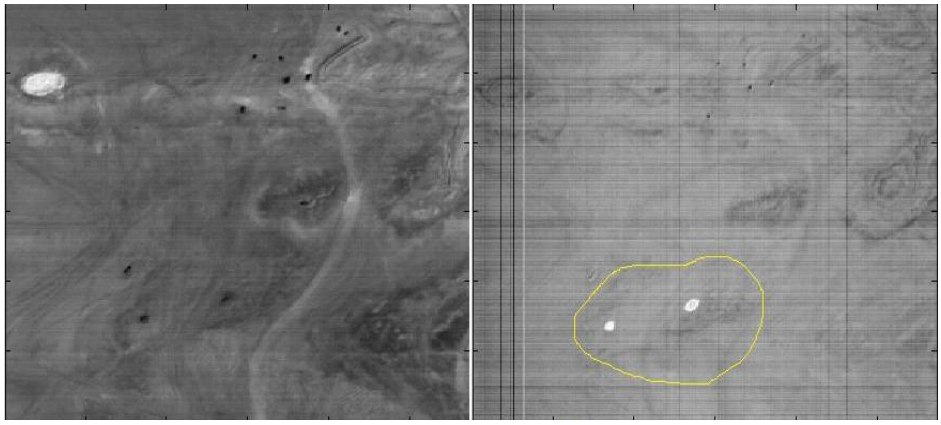

Figure 2.2: Left: Spectral lines 1400 - 1700 from the original scene. Right: The white points mark the detected targets

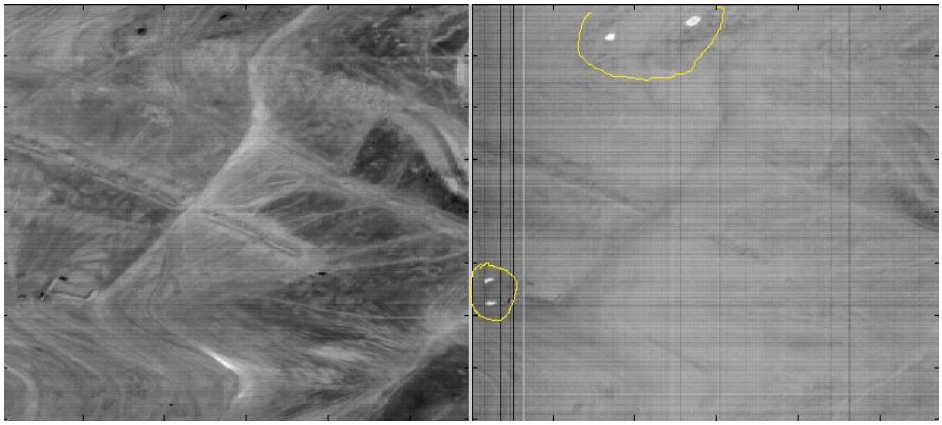

Figure 2.3: Left: Spectral lines 1600 - 1900 from the original scene. Right: The white points mark the detected targets

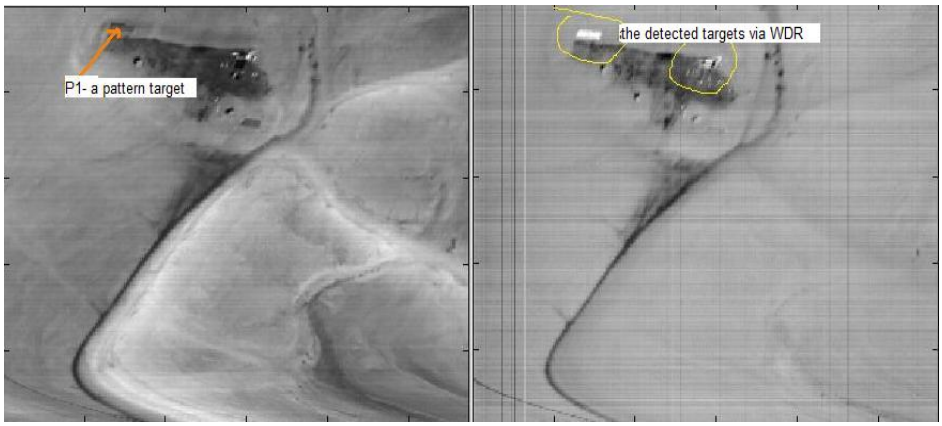

Figure 2.4: Left: Spectral lines 2200 - 2600 from the original scene. Right: The white points mark the detected targets

The detection of the suspicious points in Figs. 2.2-2.4 match exactly the known targets.

In Fig. 2.4, the point P_1 is given as a pattern of the target's material. It was extracted from the map of known targets. Its spectrum is displayed in Fig. 2.5 as a plot of the target. Other spectra plots, which were detected by the WDR algorithm in the scenes in Figs. 2.2-2.4, are classified as "spectra in suspicious points".

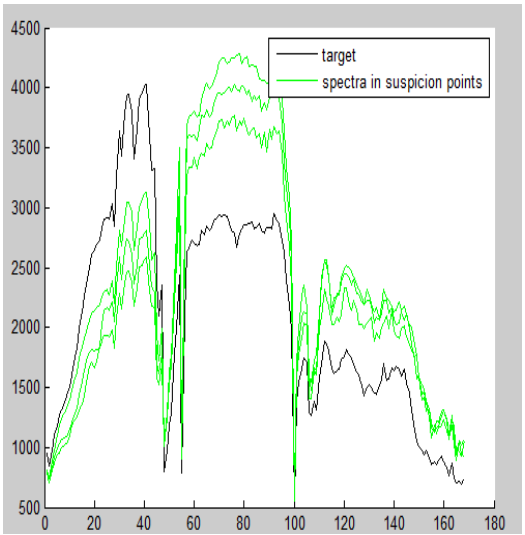

Figure 2.5: Comparison between the spectrum of the target and the spectrum from a suspicious points in Figs. 2.2-2.4. The x- and the y-axes in this figure are the wavebands and their values, respectively.

3 Method II: Unmixing by examining the neighborhood of a suspicious point (UNSP)

In this section, we provide an algorithm that detects subpixel targets. For ease of notation, a square of $m = 2m_1 + 1$ pixels on each side with a center at a pixel X is called the m-neighborhood of the pixel X. It is denoted by $\Omega_m(X)$. m_1 is the radius of this neighborhood.

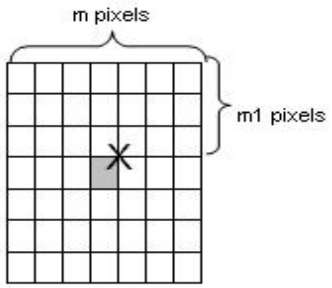

Figure 3.1: $\Omega_m(X)$ denotes the m-neighborhood of the pixel X

A connected component is a set of pixels in which any two pixels are connected to each other by paths, where a path is a sequence of pixels such that for each of its pixels the next pixel is adjacent to it either horizontally or vertically.

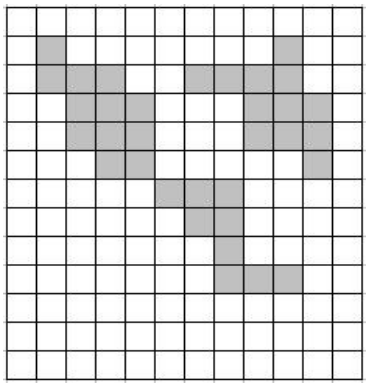

Figure 3.2: Three connected components

Consider spectra from different materials which are present in a hyperspectral image. Usually, in real situations, there is high correlation between these spectra. For example, Fig. 3.3 displays spectra of three different materials.

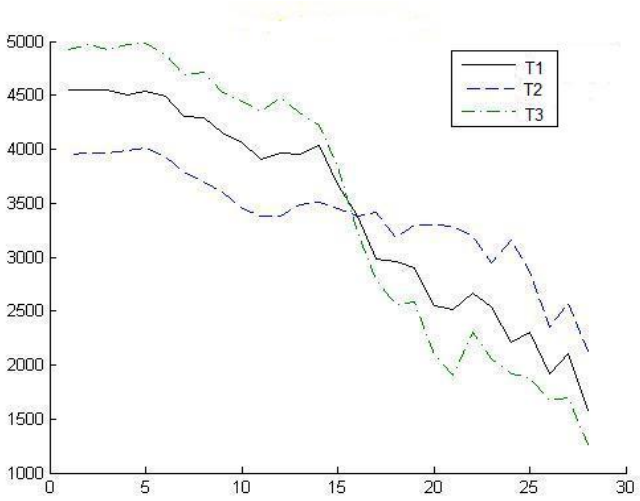

Figure 3.3: Spectra of three different materials

Each spectrum contains a linear part, which can be extracted by a least-square method (LSM). We assume that some correlation between the spectra exists because there are linear parts (trends). After the application of the first derivative to the linear part, it becomes zero and the correlation decreases. Next, we consider the first derivatives of spectra for two different materials that are less dependent.

We denote the first derivative of a spectrum of a pixel X by $d(X)$ and it is called the d-spectrum of the pixel X.

We assume that pixels, which contain target (as subpixel or as whole pixel), represent one connected component that occupies less than half of the m-neighborhood, $m = 3, 5, 7, 9$, for some pixel. T is a known given target's spectrum.

The UNSP algorithm has two steps:

Step 1. Detection of suspicious points. The following hypotheses are assumed:

H_0: Y is a suspicious point.

H_1: Y is not a suspicious point.

Hypotheses check: We consider $\Omega_m(Y)$. The indexes of these pixels are constructed. Denote a pixel located in the i-row and j-column by p_{ij}, $i, j = 1, \ldots, m$ - see Fig. 3.4. For example, $Y = p_{m_1+1 m_1+1}$ where m_1 is the radius of the neighborhood.

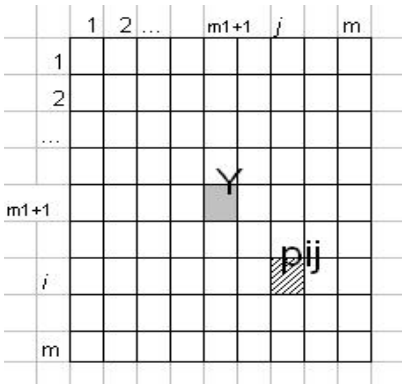

Figure 3.4: Indexing of pixels

Consider a set of pixels, denoted by Ψ, such that $corr(d(p_{ij}), d(Y)) > 0.5$, where $corr(d(p_{ij}), d(Y))$ is a correlation coefficient between the vectors $d(p_{ij})$ and $d(Y)$.

If the set Ψ either represents two or more connected components or $|\Psi| > m^2/2$, then Y is not a suspicious point. Therefore, H_1 holds. In other words, if Y is a suspicious point, then Ψ is a set of pixels that intersect with the target and this set of correlated points is concentrated around the central point Y. Here and below, we assume that a correlated point is a pixel whose d-spectrum and $d(Y)$ are correlated with correlation coefficient that is greater than 0.5.

Let N_1 be the neighborhood $\Omega_{m-2}(Y)$. N_1 is called the internal square. Let $N_2 = \Omega_m(Y) \backslash N_1$. N_2 is called the external square. They are visualized in Fig. 3.5.

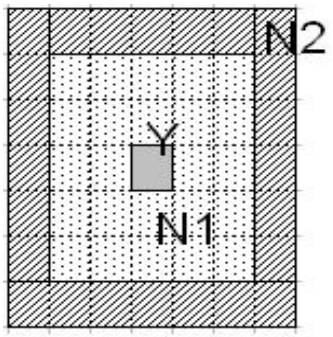

Figure 3.5: N_1 is the internal square and N_2 is the external square

Assume Δ is the set of all pixels p_{ij}, which are bounded by the external and internal squares with correlation coefficients $corr(d(p_{ij}), d(Y))$ less than 0.5 and 0.3, respectively. Each pixel in Δ is treated as a vector where its entries are all the bands (wavelengths) of this pixel. The

d-spectra of this vector is denoted by v_s where s is one of the $(i,j) \in \Delta$. The set of all these vectors is denoted by V. This is the set of all the d-spectra that belong to Δ. If $|\Delta| = s$ then $V = \{v_1, \ldots, v_s\}$.

In order to derive the d-spectrum of some material in a central pixel, the background around the central pixel has to be removed. For that, we construct an orthogonal projection ρ, which projects all the d-spectra onto the orthocomplement of the linear span where the background of the d-spectra is located. If the d-spectrum of the central pixel $d(Y)$ does not belong to this linear span, then this projection extracts an orthogonal component of $d(Y)$ which is not mixed with the background of the d-spectrum. For example, if $d(Y) = d_1 + d_2$ where d_1 belongs to the linear span generated by background of the d-spectrum and d_2 belongs to the orthocomplement of this span, then, after the projection we obtain $\rho(d(Y)) = \rho(d_2)$ which does not correlate with the background of the d-spectrum. Hence, the background's influence is removed by this projection.

Now, we formalize the above. The eigenvectors and eigenvalues of the covariance matrix E of the vectors $v_1, ..., v_s$ where $E(i,j) = v_i \cdot v_j$, $(i,j) \in \Delta$, are derived. The largest eigenvalue is denoted by max_e. The eigenvectors with eigenvalues, which are smaller than $max_e/20$ (determined experimentally), generate the eigensubspace, which is the orthocomplement of the linear span of the principal directions of the set V. Denote this orthocomplement by C.

Throughout this paper, we assume that in our model, the spectrum of any pixel X consists of three components:

1. The spectrum of a material M is different from the background;
2. The spectrum of the background constitutes from a linear combination of spectra of some pixels from the X-neighborhood;
3. Random noise is present.

The same model is true for d-spectra $P' = \tau M' + L(v_1, ..., v_s) + N$, where $P' = d(Y)$, M' is the d-spectrum of some material M, $\tau \in [0,1]$ is the portion of the material M in Y, N is random noise, $L(v_1, ..., v_s)$ is a linear combination of the vectors $v_1, ..., v_s$.

If the correlated points concentrate around Y, then these points consist of the same material as Y. If the uncorrelated points do not contain this material then they constitute the background.

Consider the orthogonal projection operator ρ. This operator projects vectors onto the orthocomplement C. The vector $\rho(L(v_1, ..., v_s))$ is approximated to be a zero vector. Thus, this orthogonal projection removes the background influence from the d-spectrum of $d(Y)$.

Let T' be the given d-spectrum of the target. If the correlation coefficient of $\rho(P')$ and $\rho(T')$ is greater than the correlation coefficient of P' and T', then Y is a suspicious point, M is a target, $T' = M'$ and H_0 holds.

Step 2. Extraction of the target spectrum from a suspicious point: Let Y be a suspicious point and T is the given target spectrum. What portion of the target is contained in Y? In Step 1, we calculated the following: V is the d-spectra set that is uncorrelated with $d(Y)$ pixels from the m-neighborhood of Y and ρ is the projection operator onto the orthocomplement of the linear span of V. Let $P_2 = \rho(d(Y))$, $T_2 = \rho(d(T))$, then $P_2 = t'T_2 + N$ where t' is an unknown parameter, N is a Gaussian random noise that is independent of T_2. The parameter $t' \in [0, 1]$ is estimated as the maximum of the independency between the two d-spectra T_2 and $P_2 - t'T_2$.

The fact that two vectors X_1 and X_2 are independent is equivalent to $corr(\varphi(X_1), \varphi(X_2)) = 0$ for any analytical function φ ([16]). An analytical function can be represented as Taylor expansion of its arguments degrees, then the condition $corr(\varphi(X_1), \varphi(X_2)) = 0$ equals to $corr((X_1)^n, (X_2)^n) = 0$ for any positive integer n where n denotes a power. In our algorithm, we limit ourself to $n = \{1, 2, 3, 4\}$. From the independency criterion of two vectors X_1 and X_2, we can take $f = |corr(X_1, X_2)| + |corr((X_1)^2, (X_2)^2)| + |corr((X_1)^3, (X_2)^3)| + |corr((X_1)^4, (X_2)^4)|$ which equals to zero in case X_1 and X_2 are independent.

If t' is estimated, then $P = t'T + B$ where P is the spectrum of the suspicious point and B is the background spectrum that is affected by noise.

3.1 Experimental results

Figures. 3.6 and 3.7 present the outputs from the application of the UNSP algorithm to two different hyperspectral sceneries.

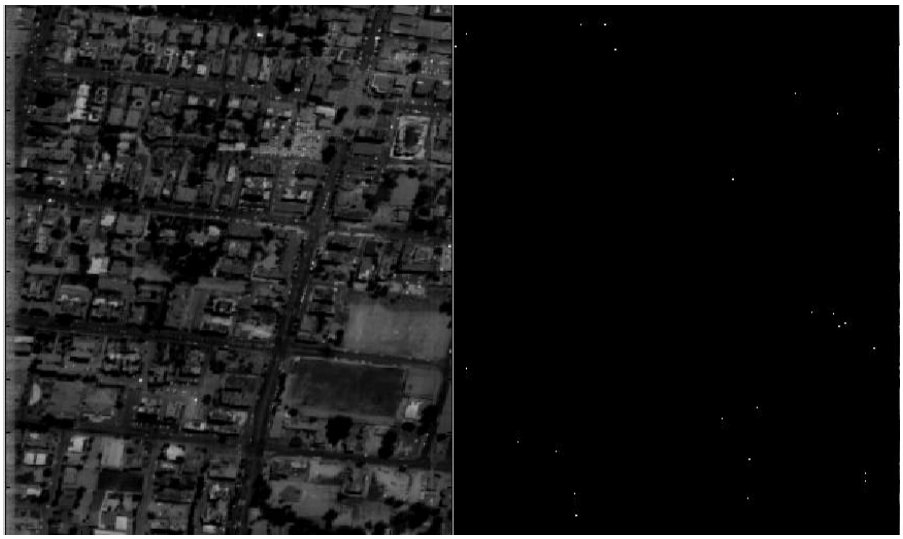

Figure 3.6: Left: The source image. Right: The white points are the suspicious points

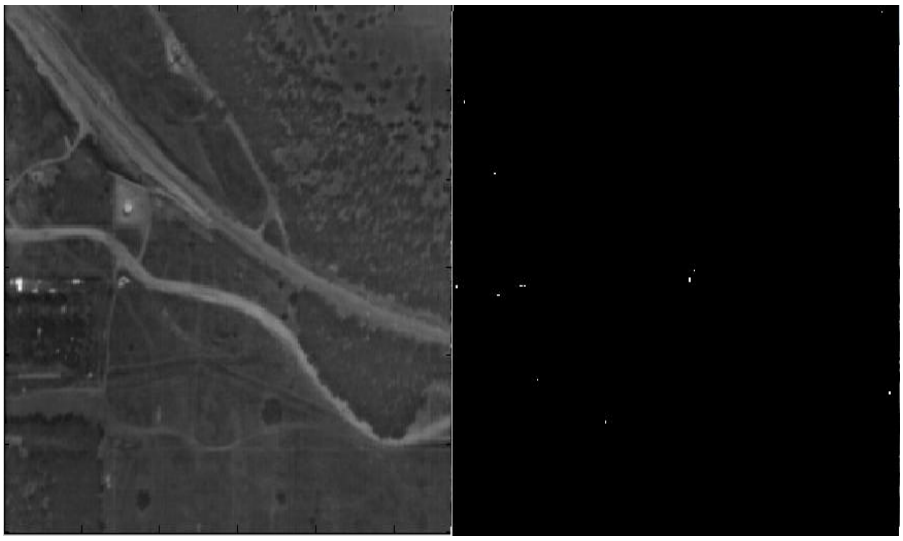

Figure 3.7: Left: The source image. Right: The white points are the suspicious points

In Figs. 3.1 and 3.1, the x- and y- axes are the wavebands and their values, respectively.

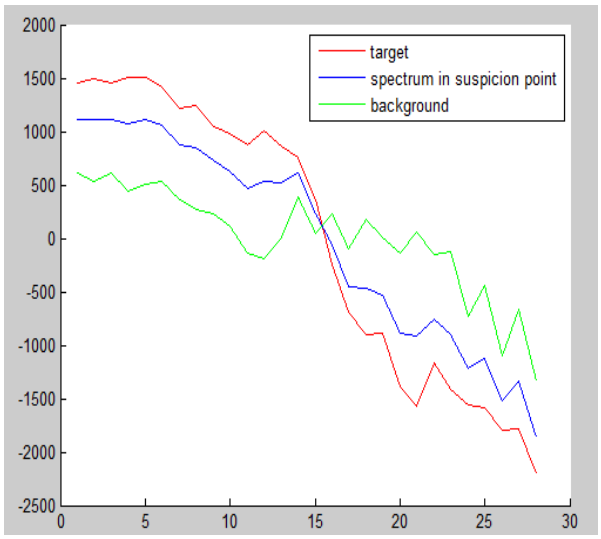

Figure 3.8: The result from the application of the UNSP unmixing algorithm to a suspicious point in Fig. 3.6. The suspicious point is decomposed into target and background spectral portions.

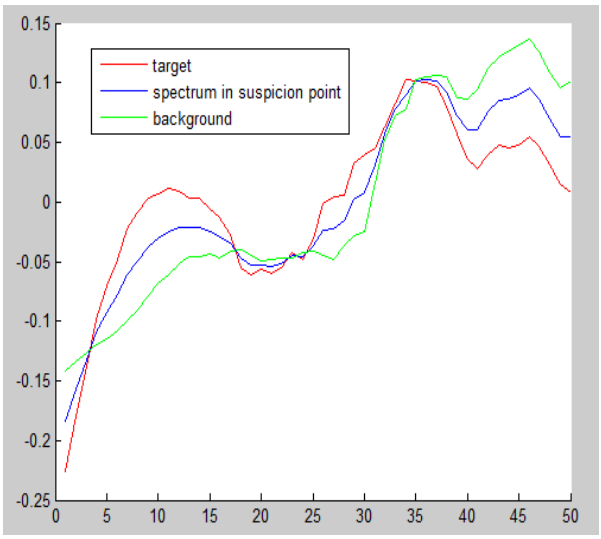

Figure 3.9: The result from the application of the UNSP unmixing algorithm to suspicious point in Fig. 3.7. The suspicious point is decomposed into target and background spectral portions.

4 Method III: Random orthogonal transformation for unmixing (ROTU)

The UNSP method, which was described in section 2 that constructed orthocomplement for the linear span of the principal directions of each pixel, is computational expansive. For independent spectra and for special type of dependent spectra, which will be defined below, we can use a faster

and less computational expansive method that is described in this section. This method works well if the $d-spectra$ of different materials are related by the so-called $sparse-independent$. This relation will be described below. If they are not $sparse-independent$, then this method generates sometimes false alarms.

Assume B is the background spectrum, P is a pixel of mixed spectrum (a spectrum of a suspicious point) and T is the given target's spectrum. For example, Figs. 4.1 and 4.2 display a scene with several suspicious points. The background, the target and the mixed pixel spectra are displayed in Fig. 4.3.

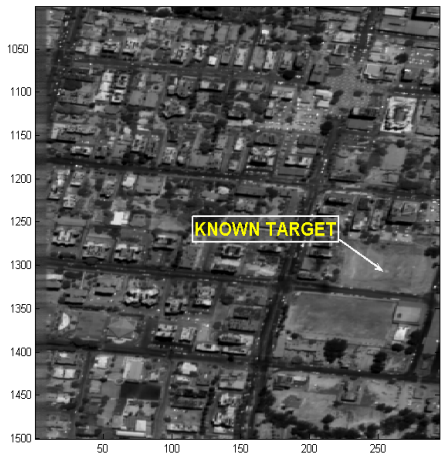

Figure 4.1: One suspicious point is pointed by the arrow.

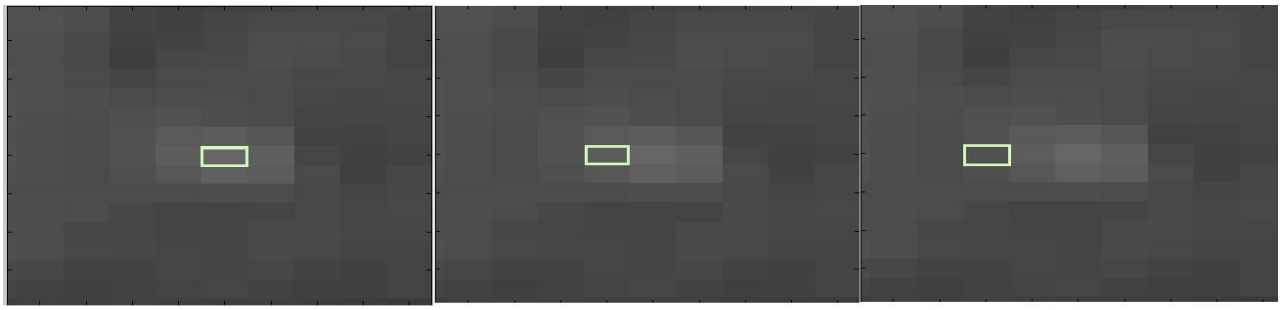

Figure 4.2: Zoom of a suspicious point and its neighborhood from Fig. 4.1. Left: The pixel with the target spectrum T. Center: The pixel with the mixed spectrum P. Right: The pixel with the background spectrum B

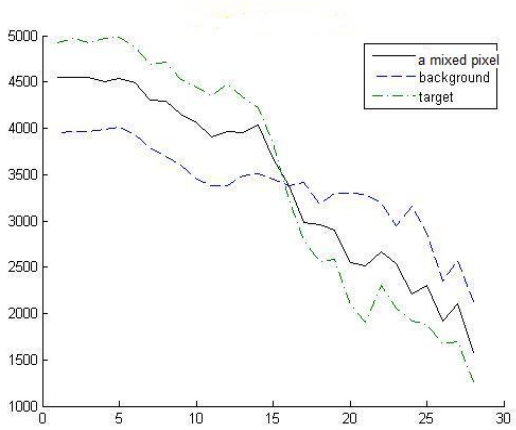

Figure 4.3: Spectra of the background, the mixed pixel and the target from Fig. 4.1

Consider three spectra: a background spectrum B, a mixed pixel spectrum (spectrum of a suspicious point) P and a target spectrum T. They are related by the following model

$$P = tB + (1-t)T, \tag{4.1}$$

which is a simplified version of Eq. 1.1, where $a_1 = t, a_2 = 1-t, s_1 = B, s_2 = T, t \in \mathbb{R}, t \in (0,1)$.

We are given the target spectrum T and the mixed pixel spectrum P. Our goal is to estimate t, denoted by $\hat{t}$, which will satisfy Eq. 4.1 provided that B and T have some independent features. Once $\hat{t}$ is found, the estimation of an unknown background spectrum B, denoted by $\hat{B}$, can be calculated as $\hat{B} = (P - \hat{t}T)/(1-\hat{t})$.

To explain our approach, we present some new definitions and preliminaries from linear algebra.

4.1 Definitions

Denote by $Pr_W(y)$ the orthogonal projection of the vector $y \in \mathbb{R}^n$ on the subspace $W \subset \mathbb{R}^n$.

Definition 4.1. *Given a vector $x = (x_1, x_2, ..., x_n)$. Let $mode_\varepsilon(x) \stackrel{\Delta}{=} argmax_p(card\{i \in [1,..,n] : |p - x_i| < \varepsilon\})$ where p is scalar. $mode_\varepsilon(x)$ is called the mode with ε-error of the vector x.*

Definition 4.2. *Let q be a positive number. A vector x is called $q - \varepsilon - sparse$ if the cardinality of its support $\{i : |x_i| > \varepsilon\}$ is less or equal to q.*

Definition 4.3. *Let L be a linear operator $\mathbb{R}^n \to \mathbb{R}^n$. The norm of the operator L is $\|L\| = \sup_{\|x\|=1} \|Lx\|$.*

Definition 4.4. *Let G be a vector space and assume that there is a finite subset $\Im \subset G$. The set $\Im$ is ε-dense if for any element $x \in G$ there is $s \in \Im$ such that $\|x - s\| < \varepsilon$.*

Definition 4.5. *Let G be a vector space and assume that the set $\Im$ is ε-dense in G. Let $V_G^r(s) = \{x \in G : \|x - s\| < r\}$. $\Im$ is homogeneous if for any $s_1, s_2 \in \Im$ and $r > \varepsilon$, $|V_G^r(s_1) \bigcap \Im| = |V_G^r(s_2) \bigcap \Im|$ holds. In this notation, $V_G^r(s)$ is called the $r-$neighborhood of s.*

For example, the set integers $\mathbb{Z}$ is ε-dense in the set of real numbers $\mathbb{R}$ when $\varepsilon = 1$. The set of rational numbers $\mathbb{Q}$ is ε-dense in $\mathbb{R}$ for any ε. $\mathbb{Z}$ and $\mathbb{Q}$ are homogeneous.

Definition 4.6. *Given two vector v_1 and v_2 in $\mathbb{R}^n$. Assume that S_1 and S_2 are two linear operators in $\mathbb{R}^n$ and L is a linear span of $\{v_1, v_2\}$. We say that S_1 and S_2 are equivalent in relation to v_1 and v_2, denoted by $S_1 \sim S_2$, if $S_1(L) = S_2(L)$.*

4.1.1 Sparse-independency definitions

Definition 4.7. *Given two vectors y_1 and y_2 where their components are denoted by $y_j(i), j = 1, 2$. Let $S_1 = \{i : |y_1(i)| > \varepsilon\}$ and $S_2 = \{i : |y_2(i)| > \varepsilon\}$. If $S_1 \bigcap S_2 = \varnothing$, then we say that y_1 and y_2 are $\varepsilon - sparse - independent$. If $\varepsilon = 0$ then they are called $sparse - independent$.*

Definition 4.8. *Let $\{e_1, e_2, ..., e_n\}$ be a basis of $\mathbb{R}^n$. Then, any subspace $W = e_{i_1}\mathbb{R} \bigoplus e_{i_2}\mathbb{R} \bigoplus \cdots \bigoplus e_{i_k}\mathbb{R}$, where $i_1, i_2, ..., i_k$ is any subset of $\{1, 2, ..., n\}$, is called a basic subspace of $\mathbb{R}^n$.*

For example, in the 3D space with the standard basis $\{e_1, e_2, e_3\}$, the plane, which is spanned by the vectors $\{e_1, e_2\}$, is a basic subspace. The plane, which is spanned by the vectors $\{(e_1 + e_3)/2, (e_2 + e_3)/2\}$, is not a basic subspace. The basic vectors $\{e_1, e_2, ..., e_n\}$ of a basis subspace must be a subset of the standard basis $\mathbb{R}^n$.

Definition 4.9. *Given a pair of vectors y_1 and y_2 in $\mathbb{R}^n$. Assume that $L(y_1, y_2)$ is the linear span of y_1 and y_2. The basic subspace $W \subset \mathbb{R}^n$ is called a dependent basic subspace of y_1 and y_2 if $dim(Pr_W L(y_1, y_2)) = 1$.*

Definition 4.10. *Given a pair of vectors y_1 and y_2 in $\mathbb{R}^n$. A natural number r is called a basis dependent rank for the vectors y_1 and y_2, if the two conditions hold:*

1. *If W is a dependent basic subspace of y_1 and y_2, then $dim(W) \leq r$;*
2. *There exists a dependent basic subspace W of y_1 and y_2 where $dim(W) = r$.*

For example, assume that $y_1 = (1, 2, 2, 3)$ and $y_2 = (2, 1, 1, -1)$ belong to $\mathbb{R}^4 = e_1\mathbb{R} \bigoplus \cdots \bigoplus e_4\mathbb{R}$. Let $W = e_2\mathbb{R} \bigoplus e_3\mathbb{R}$. Then, $z_1 = Pr_W(y_1) = (2, 2)$, $z_2 = Pr_W(y_2) = (1, 1)$ and $z_1 = 2z_2$. Thus, we obtain that W is a dependent basic subspace for y_1 and y_2. It is clear that the dependency rank of the vectors y_1 and y_2 is 2.

If a pair of vectors y_1 and y_2 are sparse-independent, then their basis dependent rank is zero.

Assume that a vector $y \in \mathbb{R}^n$. The set $S_\varepsilon(y) = \{i : |y(i)| > \varepsilon\}$ is called the ε-support of y.

Definition 4.11. *Given a pair of vectors y_1 and y_2 in $\mathbb{R}^n$ and $\varepsilon > 0$. Let*

$$r = min[card(S_{\varepsilon^2}(y_1) \backslash (S_\varepsilon(y_1) \bigcap S_\varepsilon(y_2)), card(S_{\varepsilon^2}(y_2) \backslash (S_\varepsilon(y_1) \bigcap S_\varepsilon(y_2))].$$

r is called the ε-sparse-independent-rank of the vectors y_1 and y_2

Definition 4.12. *Given a pair of vectors y_1 and y_2. Assume that r is their ε-sparse-independent-rank and b is a basis dependent rank of the vectors y_1 and y_2. These two vectors are called a ε-sparse-ergodic independent if $r > b$.*

For example, sparse-independent vectors have a basis dependent rank zero and they are ε-sparse-ergodic independent.

4.2 Preliminaries

The goal of this section is to produce the conditions for which vectors provide a solution for the unmixing problem given by Eq. 4.1.

Lemma 4.1. *Let B be an unknown background spectrum, T and P are known spectra of the target and of the mixed pixel, respectively. The parameter t satisfies the linear mixing model $P = tT + (1-t)B$. If for any fixed positive value $\gamma > 0$, $0 < \varepsilon < 1$, $S_1 = \{i : |B(i)| > \gamma\}$ and $S_2 = \{i : |T(i)| > \gamma/\varepsilon\}$ we have that $card(S_2 \setminus S_1) \geq card(S_1 \cap S_2)$, then $t = mode_\varepsilon \zeta$ (definition 4.1) where ζ is a vector whose coordinates are $\{\frac{P}{T_1}\}$ and $T_1 = \{T | \|T\| > \gamma/\varepsilon\}$.*

Proof: Let $\xi = \frac{P}{T_1} = \frac{tT+(1-t)B}{T_1} = t + (1-t)\frac{B}{T_1}$ and let $\zeta = \frac{B}{T_1}$, then $\xi = t + (1-t)\zeta$. Let $W = (S_2 \setminus S_1)$ and $V = (S_2 \cap S_1)$, then $|W| \geq |V|$, $W \bigcap V = \varnothing$, $\{\zeta(i)\}_{i \in W} \bigcap \{\zeta(i)\}_{i \in V} = \varnothing$ and $\{\zeta(i)\} = \{\zeta(i)\}_{i \in W} \bigcup \{\zeta(i)\}_{i \in V}$. If $i \in W$ then $|\zeta(i)| = \frac{B(i)}{T(i)} < \gamma/(\gamma/\varepsilon) = \varepsilon$ and $card(\{i | |\zeta(i)| \geq \varepsilon\}) \leq |V| \leq |W|$. Thus, $mode_\varepsilon(\zeta) = 0$ and $mode_\varepsilon(\xi) = t$. ■

From Lemma 4.1 we get the following proposition.

Proposition 4.1. *Let B be an unknown background spectrum, T and P are known spectra of the target and of the mixed pixel, respectively. The parameter t satisfies the linear mixing model $P = tT + (1-t)B$. Let $T_1 = \{T | \|T\| > \gamma\}$ and ξ is a vector with the coordinates $\{\frac{P}{T_1}\}$. Assume that Let B and P are $\varepsilon - sparce - independent$. Then, $mode_{\varepsilon/\gamma}(\xi) = t$.*

Proposition 4.2. *Assume that B is unknown background spectrum where T and P are the known spectra of the target and of the mixed pixel, respectively. Assume that the parameter t satisfies the linear mixing model $P = tT + (1-t)B$. If B and T are orthogonal, then $t = \langle T, P \rangle / \|T\|^2$.*

Proof: $\langle T, P\rangle = \langle T, t\cdot T + (1-t)B\rangle = t\cdot\langle T, T\rangle = t\cdot\|T\|^2$ ■

Note that if $\langle T, B\rangle = \epsilon$, then the variance of this estimation's error is $\epsilon/\|T\|^2$. The "sparse-independency" and orthogonality are very strong conditions. The next theorem shows that "ε-sparse-ergodic independency" also provides a solution for the linear unmixing problem.

Theorem 4.1. *Given $\varepsilon > 0$, B is an unknown background spectrum, T and P are known spectra of the target and of the mixed pixel, respectively. The parameter t satisfies the linear mixing model $P = tT + (1-t)B$. Assume B and P are ε-sparse-ergodic independent. T_i denotes the ith coordinate of the vector T. Let $T_1 = \{T_i : |T_i| > \varepsilon\}$ and ξ is a vector with the coordinates $\{\frac{P}{T_1}\}$. Then, $mode_\varepsilon(\xi) = t$.*

Proof: Let $\xi = \{\frac{P}{T_1}\}$ and $\zeta = \{\frac{B}{T_1}\}$. $\xi = t + (1-t)\zeta$, then we only need to prove that $mode_\varepsilon(\zeta) = 0$. Assume that $p \neq 0$ is any real number. Let $\Omega = \{i | B_i = pT_i\}$ and $\Lambda = \{i | B_i < \varepsilon^2 \wedge T_i > \varepsilon\}$. The ε-sparse-ergodic independency of B and T follows from the fact that $card(\Omega) < card(\Lambda)$, because if $W = \bigoplus_{i\in\Omega} e_i\mathbb{R}$, then W is a dependent basic subspace of T and B and $dim(W) = card(\Omega)$. Note that $\{i | B_i < \varepsilon^2 \wedge T_i > \varepsilon\} = \{i | \zeta_i < \varepsilon\}$ and $\{i | B_i = pT_i\} = \{i | \zeta_i = p\}$. From the condition $card(\{i | \zeta_i = p\}) < card(\{i | \zeta_i < \varepsilon\})$ follows that $mode_\varepsilon(\zeta) = 0$ which means that $mode_\varepsilon(\xi) = t$. ■

Now we show how to obtain the "ε-sparse-ergodic independency".

Theorem 4.2. *Assume that $\{v_1, v_2\} \subset \mathbb{R}^n$, $\langle v_1, v_2\rangle = 0$, $\varepsilon > 0$. Let $\Lambda = \{S_1, S_2, ..., S_r\}$ be the ε-dense and homogeneous set of isometrics operators in $\mathbb{R}^n$. Then,*

$w_1 \triangleq [v_1, S_1(v_1), S_2(v_1), ..., S_r(v_1)]$ and $w_2 \triangleq [v_2, S_1(v_2), S_2(v_2), ..., S_r(v_2)]$ are ε-sparse-ergodic independent.

The rest of this section proves Theorem 4.2

Proposition 4.3.

For any pair of orthogonal vectors $\{v_1, v_2\} \subset \mathbb{R}^2$ such that $\langle v_1, v_2\rangle = 0$, for any linear operator S to one-dimensional subspace with $\|S\| = 1$ and for any real $t \neq 0$, the inequality $P[S(v_2) = t\cdot S(v_1)] < P[S(v_2) = 0]$ holds where P is probability.

Proof: Any linear operator $S : \mathbb{R}^2 \to \mathbb{R}^1$ with $\|S\| = 1$, can be represented as $S(x) = \langle s, x\rangle$ where $s = (s_1, s_2)$ is a vector with $\|s\| = 1$.

Consider

$$\hat{S} = \begin{pmatrix} s_1 & s_2 \\ -s_2 & s_1 \end{pmatrix} = \begin{pmatrix} cos(\varphi) & sin(\varphi) \\ -sin(\varphi) & cos(\varphi) \end{pmatrix}$$

to be a rotation of $\mathbb{R}^2$. This rotation splits the linear operator S into a product of the rotation operator with the orthogonal projection such that $S = \hat{S} \circ Pr_1$ where $Pr_1(x_1, x_2) \triangleq (x_1, 0)$.

We can choose a basis $\{e_1, e_2\}$ of $\mathbb{R}^2$ such that $v_2 = (a, 0)$, $v_1 = (0, b)$. In this representation, $\frac{S(v_2)}{S(v_1)} = -\frac{a \cdot sin(\varphi)}{b \cdot cos(\varphi)} = (-a/b)\tan(\varphi)$. For any $t \in \mathbb{R}$, the solution of $t = (-a/b)\tan(\varphi)$ exists.

In addition, $\tan'(\varphi) = \frac{1}{\cos^2(\varphi)}$ has a minimum in zero. Hence, the size of the set $\{\varphi : |\tan(\varphi)| < \varepsilon\}$ is more than the size of the set $\{\varphi : |\tan(\varphi) - t| < \varepsilon\}$ for $t \neq 0$. ■

Corollary 4.1. *Let $\{S_1, S_2..., S_r\}$ be ε-dense and homogeneous set (definitions 4.4 and 4.5) of the linear operators $\mathbb{R}^2 \to \mathbb{R}^1$ and $\{v_1, v_2\}$ is orthogonal in $\mathbb{R}^2$. Then, the vector ς with the coordinates $\varsigma_i = \frac{S_i(v_1)}{S_i(v_2)}, \quad i = 1, \ldots, r$, has $mode_\varepsilon(\varsigma) = 0$.*

Proposition 4.4. *Given a pair of orthogonal vectors $\{v_1, v_2\} \subset \mathbb{R}^n$, $\langle v_1, v_2 \rangle = 0$ and $\varepsilon > 0$. Assume that $S_1, S_2, ..., S_r$ is an ε-dense and homogeneous set of isometrics operators in $\mathbb{R}^n$. Then, the vector ς, which has the coordinates $\varsigma_{i,\alpha} = \chi_i^\alpha, \ i = 1, \ldots, n, \ \alpha = 1, \ldots, r$, has $mode_\varepsilon(\varsigma_{i,\alpha}) = 0$ where χ_i^α is the ith coordinate of the vector $\chi^\alpha = \left(\frac{(S_\alpha(v_1))_1}{(S_\alpha(v_2))_1}, ..., \frac{(S_\alpha(v_1))_n}{(S_\alpha(v_2))_n}\right)$.*

Proof: Denote $\Lambda = \{S_1, S_2, ..., S_r\}$. The set of χ coordinates is divided into the union of the subsets τ_i where $\tau_i = \{\varsigma_{i,\alpha}\}_{S_\alpha \in \Lambda}$, $\alpha = 1, \ldots, r$.

We have to show that for $mode_\varepsilon(\tau_i) = 0$, $i = 1, \ldots, n$. Assume that θ_i is a projection to coordinate i, such that for $x = (x_1, ..., x_n)$, $\theta_i x = x_i$. Assume that L is a linear span of $\{v_1, v_2\}$. Define $R_\alpha = \theta_i \circ S_\alpha$ where $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^1$. Our assumption was that $\Lambda = \{S_1, S_2, ..., S_r\}$ is ε-dense and homogeneous. Therefore, $\{R_1, R_2, ..., R_r\}$ is ε-dense and homogeneous too. Then, Corollary 4.1 implies $mode_\varepsilon(\tau_i) = 0$.

■

Proof of Theorem 4.2.

Let L be the linear span of w_1 and w_2 from $\mathbb{R}^{nr} = \bigoplus_{S \subset \Lambda} \mathbb{R}^n$. If the theorem is not true, then there is a basic subspace W in $\bigoplus_{S \subset \Lambda} \mathbb{R}^n$ such that $dim(W) = q$ and $dim(Pr_W(L)) = 1$. It means that $w_1|_W = s$ and $w_2|_W = t \cdot s$, $t \neq 0$, such that $dim(W) > card\{i | (w_2)(i) = 0 \text{ and } (w_1)(i) \neq 0\}$. This means that a vector with the coordinates $\varsigma_i = \frac{(w_2)(i)}{(w_1)(i)}, \quad i = 1, \ldots, nr$, has $mode_\varepsilon(\varsigma_i) = t \neq 0$. This contradictes Proposition 4.4.

4.3 The unmixing algorithm

Assume that B is a background spectrum, P is a pixel of mixed spectrum (a spectrum of a suspicious point) and T is a given target's spectrum. Assume that they are related by the linear mixing model $P = tT + (1 - t)B$. We want to estimate the parameter t.

In general, spectra of different materials have different sections: some sections are mutually correlated, some sections are smooth, some sections are not random, some have sections with oscillatory behavior. First and second large derivatives are used as characteristic features for spectral classification.

We assume that the spectra are autocorrelated. Let $D = (D_1, D_2)$ be a pair of first and second derivatives operator. Our assumptions imply that $D(T)$ and $D(B)$ are more independent than T and B. The D operator reduces the correlation between spectra and it makes them more independent.

Let x be a spectrum. We denote by a $d-spectrum$ the vector Dx. If two $d-spectra$ of different materials are statistically independent (or orthogonal), then estimating t can be obtained by using Eq. 4.2. Statistical independency is a very strong condition. Sometimes we have $sparse-independency$ for the $d-spectra$. The sparse-independency condition yields an estimation from Proposition 4.1. In general, we have ε deviation from orthogonality and the condition of $sparse-independency$ is never true.

The unmixing algorithm has the following steps:

Step 1: The operator D is applied to T and P. We assume that $|corr(DT, DB)| < \varepsilon$.

Step 2: Assume that $\Lambda = \{S_1, S_2, ..., S_r\}$ is random isometric rotations in $\mathbb{R}^n$, $W : \mathbb{R}^n \rightarrowtail \mathbb{R}^{2nr} = \bigoplus_{S\subset\Lambda} \mathbb{R}^n$ is an embedding of $\mathbb{R}^n$ into $\mathbb{R}^{2rn}$, and

$$w_T \triangleq WT = [DT, S_1DT, S_2DT, ..., S_rDT], \tag{4.2}$$

$$w_B \triangleq WB = [DB, S_1DB, S_2DB, ..., S_rDB], \tag{4.3}$$

$$w_P \triangleq WP = [DP, S_1DP, S_2DP, ..., S_rDP]. \tag{4.4}$$

Step 3: Construct the vector ς with the coordinates

$$\varsigma_{\alpha,i} = \frac{(w_P)_i}{(w_T)_i}, \quad \alpha = 1, \ldots, r, \quad i = 1, \ldots, 2n. \tag{4.5}$$

Step 4: Estimation of t, denoted by $\hat{t}$, is computed by $\hat{t} = mode_\varepsilon \left(\frac{(w_P)_i}{(w_T)_i}\right), i = 1, \ldots, 2rn$. B is estimated by $\hat{B} = (P - \hat{t}T)/(1 - \hat{t})$.

The vectors w_T, w_B and w_P are related by $w_P = t \cdot w_T + (1 - t) \cdot w_B$, where w_B is unknown. Proposition 4.4 and Theorems 4.1 and 4.2 yield that $\{w_T, w_B\}$ are ε-sparse-ergodic independent and the vector ς with the coordinates

$$\varsigma_{\alpha,i} = \frac{(S_\alpha D_1 B)_i}{(S_\alpha D_1 T)_i}, \quad \alpha = 1, \ldots, r, \quad i = 1, \ldots, 2n \tag{4.6}$$

has $mode_\varepsilon(\varsigma) = 0$.

It follows, that $mode_\varepsilon(\{\frac{(w_P)_i}{(w_T)_i}\}_{i=1}^{2rn})$ is an estimation of t. As we increase the number of operators then the estimate becomes more accurate.

Definition 4.13. *Equations 4.2-4.6 denote a vector ζ with the set of coordinates $\zeta_k = \frac{(w_P)_k}{(w_T)_k}$ $k = 1, \ldots, 2rn$. Assume $\varepsilon > 0$. Assume $b = \{0, \varepsilon, 2\varepsilon, 3\varepsilon, 4\varepsilon, \ldots, 1 = m\varepsilon\}$ is a partition of $[0, 1]$. Then, the vector H such that $H_i = card\{k | i\varepsilon - \varepsilon/2 < \zeta_k \leq i\varepsilon + \varepsilon/2\}$, $i = 1, \ldots, m$ is an histogram of ζ_k related to partition b. This histogram is called the "unmixing histogram".*

Figure 4.4 is the "unmixing histogram" for B and P from Fig. 4.3.

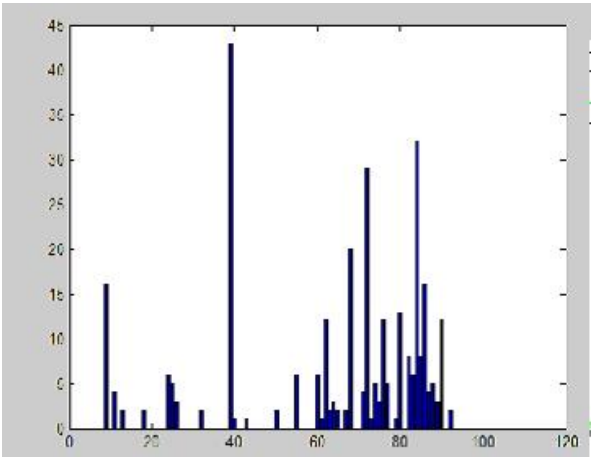

Figure 4.4: The "unmixing histogram" for Fig. 4.3. The x axis is the partition b of $[0, 1]$. The y axis is the number of coordinate that belong to each interval in the b partition. $\varepsilon = 0.01$.

We can see from Fig. 4.4 the distribution of the values of vector ζ on the interval $[0, 1]$ where the partition step is 0.01. We can see from its histogram that the mode of this distribution is $0.4 = 40 \cdot 0.01$. t is estimated by $\hat{t} = mode_\varepsilon(\xi) = 0.4$, where the vector ξ is taken from Eq. 4.6. From $\hat{t}$ we get the estimation of the background $\hat{B} = (P - \hat{t}T)/(1 - \hat{t})$. The comparison between B and $\hat{B}$ is shown in Fig. 4.5.

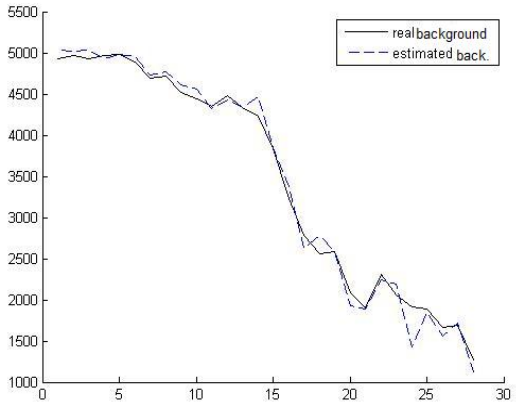

Figure 4.5: The real and the estimated spectrum of the background.

4.4 Experimental results

1. Experiment with random spectra signatures. Assume the vectors B and T were randomly generated. B is unknown. The vector P was computed by $P = tT + (1-t)B$ where $t = 0.1, 0.3, 0.7, 0.9$. Figure 4.6 displays the plots of P, T and B for $t = 0.7$. Vector P corresponds to spectrum of a mixed pixel, T corresponds to a known target spectrum and B corresponds to an unknown background spectrum.

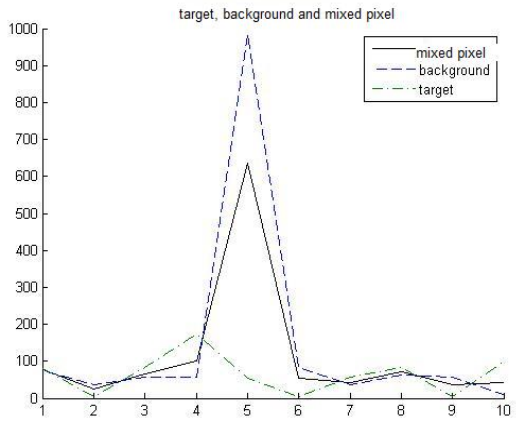

Figure 4.6: The random vector signatures.

The ROTU algorithm is applied to the pair P and T with different t values. The ROTU algorithm produces an estimate of t. Then, the estimate of $\hat{t}$ and t are compared. Once $\hat{t}$ is obtained then the unknown B is estimated by $\hat{B} = (P - \hat{t}T)/(1-\hat{t})$.

The "unmixing histogram" from definition 4.13 is displayed in Fig. 4.7.

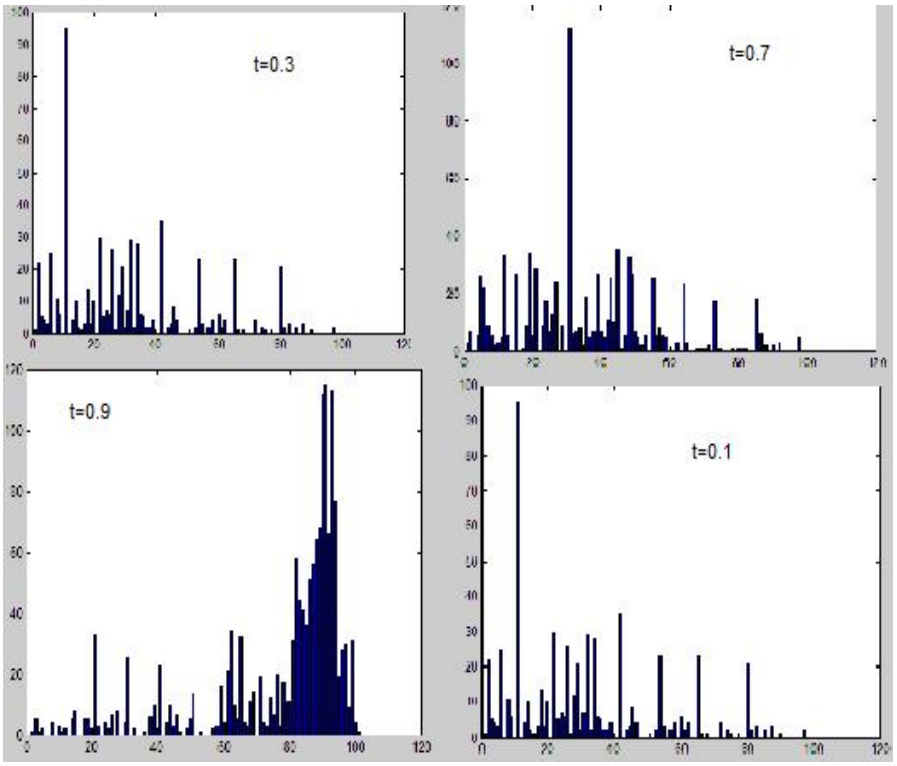

Figure 4.7: The "unmixing histogram" from definition 4.13 for different values of t. The horizontal and the vertical axes mean a partition b of $[0, 1]$. The cardinality of the coordinates set belongs to each interval in the partition b. $\varepsilon = 0.01$.

Estimations of the true values of t by $\hat{t}$ are:

$t = 0.3$ vs. $\hat{t} = 0.3$, $t = 0.7$ vs. $\hat{t} = 0.7$, $t = 0.9$ vs. $\hat{t} = 0.9$ and $t = 0.1$ vs. $\hat{t} = 0.1$. In all these cases, $\hat{t} = t$.

2. **Experiment with spectra from real materials.** The spectra B and T of two materials were taken from a database. They are represented in Fig. 4.8. P is simulated by $P = tT + (1 - t)B$ for $t = 0.3, 0.55, 0.7, 0.8$.

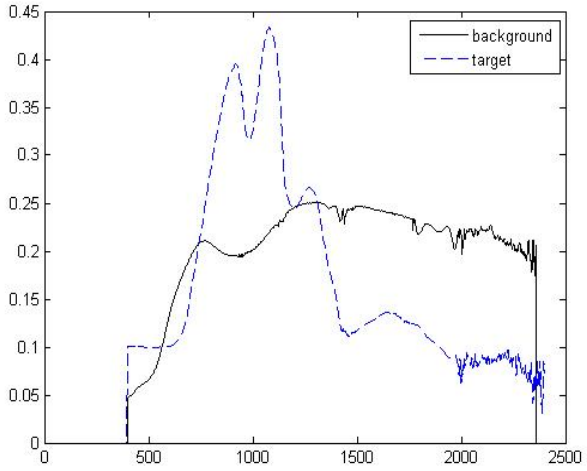

Figure 4.8: The materials T and B where the x- and y-axes are the wavebands and their values, respectively.

P and T are known while B is unknown. The ROTU algorithm is applied to the pair P and T for each value t. The ROTU algorithm estimates t by $\hat{t}$. Once $\hat{t}$ is obtained, then the unknown B can be estimated by $\hat{B} = (P - \hat{t}T)/(1 - \hat{t})$.

The "unmixing histograms" (definition 4.13) and the comparison between B and its estimate $\hat{B}$ are given below.

In each of the Figs. 4.9-4.12, the horizontal axes of the left images represent the partition b of $[0, 1]$. The vertical axes of the left images represent the cardinality of the coordinate's set that belong to each interval in the partition of b. $\varepsilon = 0.01$.

Case 1. t=0.3; $\hat{t}$=0.28.

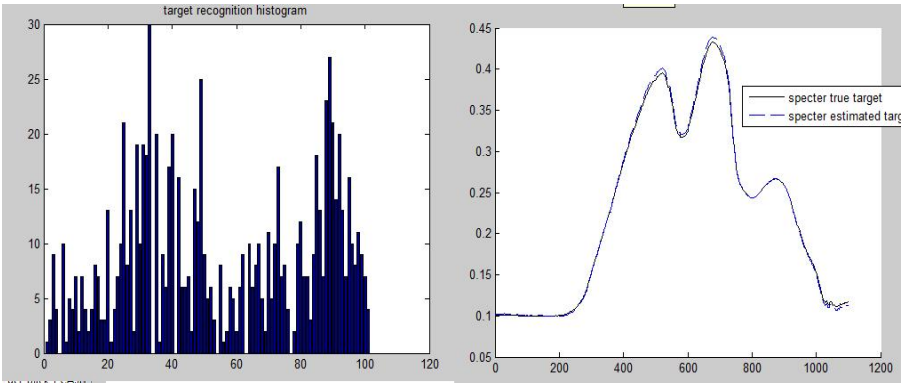

Figure 4.9: Left: The "unmixing histogram" (definition 4.13). Right: Comparison between T and its estimate $\hat{T}$.

Case 2. t=0.7; $\hat{t}$=0.69.

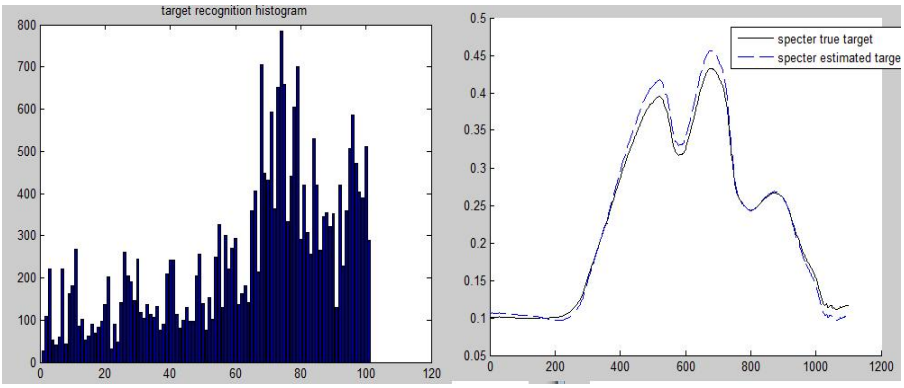

Figure 4.10: Left: The "unmixing histogram" (definition 4.13). Right: Comparison between B and its estimate $\hat{B}$.

Case 3. t=0.55; $\hat{t}$=0.59.

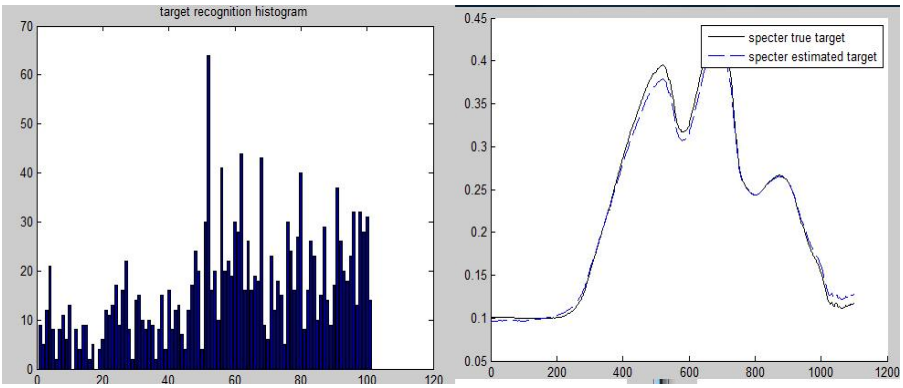

Figure 4.11: Left: The "unmixing histogram" (definition 4.13). Right: Comparison between B and its estimate $\hat{B}$.

Case 4. t=0.8; $\hat{t}$=0.82.

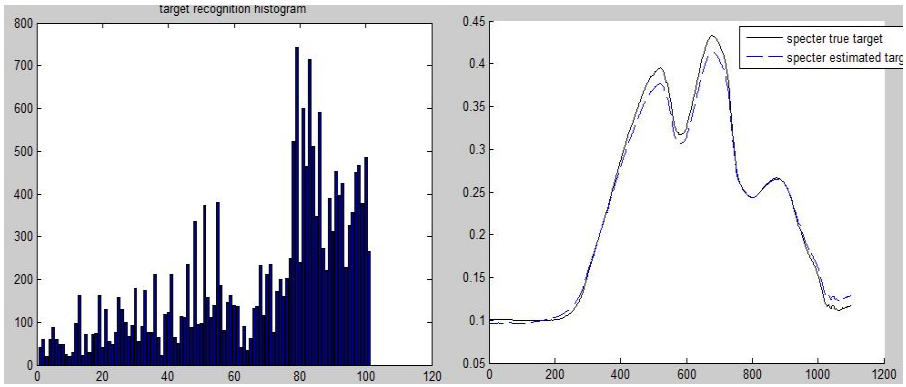

Figure 4.12: Left: The "unmixing histogram" 4.13. Right: Comparison between B and its estimate $\hat{B}$.

3. Experiment with subpixel's target recognition. We present the results after the application of the ROTU method to two different scenes. One scene has targets that occupy more than a pixel. The other scene has a subpixel target. In each figure, the parameter t, which was introduced in 4.1, means portion of the target's material in the detected pixel.

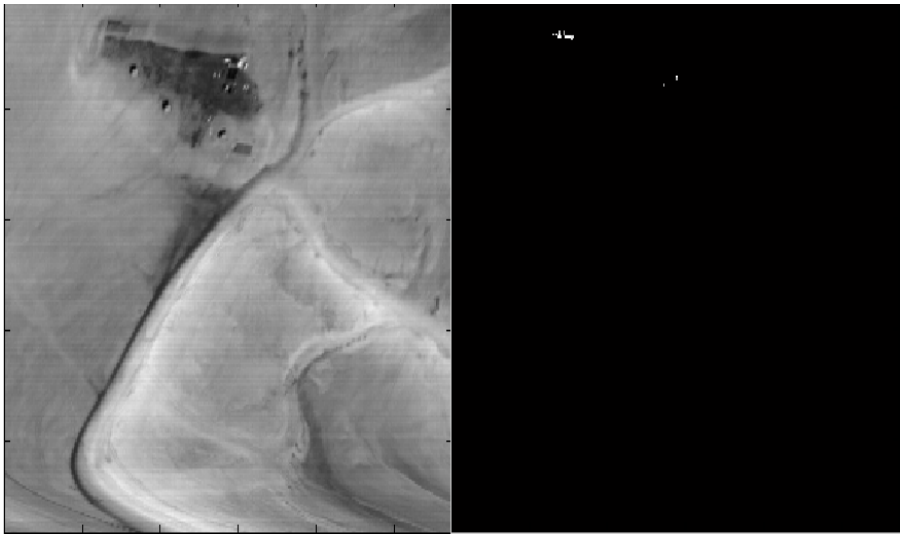

Figure 4.13: Left: The scene. Right: Points that contain a target when $t = 1$.

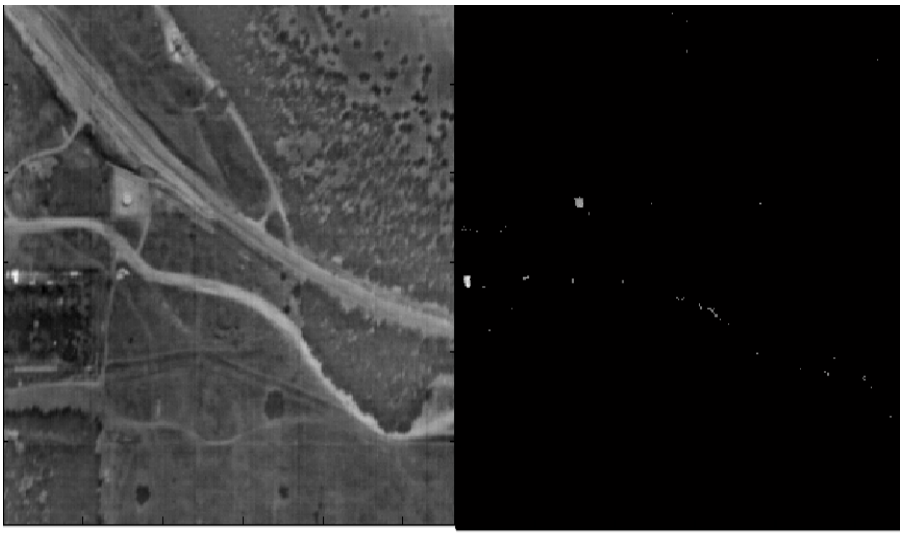

Figure 4.14: Left: The scene. Right: points that contain a target when $t > 0.5$.

Conclusions

We presented three algorithms for linear unmixing. The first algorithm (titled WDR) works well but does not detect sub-pixel targets. The second algorithm (titled UNSP) works well but it is computational expensive due to the need to search in each pixel's neighborhood. The third algorithm (titled ROTU) less reliable as WDR and UNSP but works much faster. It generates more false alarms.

In the future, we plan to add to these algorithms a classification method with machine learning methodologies to separate between the background spectra and the target's spectrum.

References

[1] A. Ifarraguerri and C.-I Chang, *Multispectral and hyperspectral image analysis with convex cones,* IEEE Trans. Geosci. Remote Sens., vol. 37, no. 2, pp. 756770, Mar. 1999.

[2] J. Boardman, *Automating spectral unmixing of AVIRIS data using convex geometry concepts,* in Summaries 4th Annu. JPL Airborne Geoscience Workshop, vol. 1, 1993, JPL Pub. 93-26, pp. 1114.

[3] M. D. Craig, *Minimum-volume transforms for remotely sensed data,* IEEE Trans. Geosci. Remote Sens., vol. 32, no. 1, pp. 99109, Jan. 1994.

[4] C. Bateson, G. Asner, and C. Wessman, *Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis,* IEEE Trans. Geosci. Remote Sens., vol. 38, no. 2, pp. 10831094, Mar. 2000.

[5] M. E. Winter, *N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data,* in Proc. SPIE Conf. Imaging Spectrometry V, 1999, pp. 266275.

[6] José M. P. Nascimento and José M. Bioucas-Dias. Vertex component analysis: A fast algorithm to unmix hyperspectral data. IEEE Trans. Geosci. Remote Sensing, 43(4):898–910, 2005.

[7] Chein-I Chang. *Hyperspectral Imaging: Techniques for spectral detection and classification.* Kluwer Academic, New York, 2003.

[8] Chein-I Chang, X. Zhao, M. L. G. Althouse, and J. J. Pan. *Least squares subspace projection approach to mixed pixel classification for hyperspectral images.* IEEE Trans. Geosci. Remote Sensing, 36(3):898–912, 1998.

[9] J. C. Harsanyi and Chein-I Chang. *Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach.* IEEE Trans. Geosci. Remote Sensing, 32(4):779–785, 1994.

[10] Y. H. Hu, H. B. Lee, and F. L. Scarpace. *Optimal linear spectral unmixing.* IEEE Trans. Geosci. Remote Sensing, 37:639–644, 1999.

[11] L. O. Jimenez and D. A. Landgrebe. *Hyperspectral data analysis and supervised feature reduction via projection pursuit.* IEEE Trans. Geosci. Remote Sensing, 37(6):2653–2664, 1999.

[12] A. S. Mazer, M. Martin, et al. Image processing software for imaging spectrometry data analysis. Rem. Sens. of the Environ., 24(1):201–210, 1988.

[13] J. J. Settle. *On the relationship between spectral unmixing and subspace projection.* IEEE Trans. Geosci. Remote Sensing, 34:1045–1046, 1996.

[14] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman. *Discrimination among semi-arid landscape endmembres using the spectral angle mapper (SAM) algorithm.* In Summaries of the 3rd annu. JPL Airborne Geosci. Workshop, R. O. Green, Ed. Publ., 92-14, volume 1, pages 147–149, 1992.

[15] P. Common. *Independent component analysis: A new concept.* Signal Processing, 36:287–314, 1994.

[16] Aapo Hyvarinen, J. Karhunen, and E. Oja. *Independent Component Analysis.* John Wiley & Sons, Inc., 2001.

[17] Jessica D. Bayliss, J. Anthony Gualtieri, and Robert F. Cromp. *Analysing hyperspectral data with independent component analysis.* In Proc. of the SPIE conference 26th AIPR Workshop: Exploiting New Image Sources and Sensors, volume 3240, pages 133–143, 1997.

[18] V. Botchko, E. Berina, Z. Korotkaya, J. Parkkinen, and T. Jaaskelainen. *Independent component analisys in spectral images.* In Proc. of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation, pages 203–207, 2003.

[19] Nirmal Keshava, Jonh Kerekes, Dimitris Manolakis, and Gary Shaw. *An algorithm taxonomy for hyperspectral unmixing.* In Proc. of the SPIE AeroSense Conference on Algorithms for Multispectral and Hyperspectral Imagery VI, volume 4049, pages 42–63, 2000.

[20] Lucas Parra, Klaus-Robert Mueller, Clay Spence, Andreas Ziehe, and Paul Sajda. *Unmixing hyperspectral data.* Advances in Neural Information Processing Systems, 12:942–948, 2000.

[21] Te Ming Tu. *Unsupervised signature extraction and separation in hyperspectral images: A noise-adjusted fast independent component analysis approach.* Optical Engineering of SPIE, 39(4):897–906, 2000.

[22] H. Attias. *Independent factor analysis.* Neural Computation, 11(4):803–851, 1999.

[23] Eric Moulines, Jean-Francois Cardoso, and Elisabeth Gassiat. *Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models.* In Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, volume 5, pages 3617–3620, 1997.

[24] José M. P. Nascimento and José M. Bioucas-Dias. *Does independent component analysis play a role in unmixing hyperspectral data?* IEEE Trans. Geosci. Remote Sensing, 43(1):175–187, 2005.

[25] José M. P. Nascimento and José M. Bioucas-Dias. *Dependent Component Analysis: A Hyperspectral Unmixing Algorithm* , in Proceedings of the 3rd IbPRIA, ser. LNCS, vol. 4478. Springer-Verlag, pp. 612–619, Girona, Spain, June 2007.

[26] C.H. Chen and X. Zhang. *Independent component analysis for remote sensing study.* In Proc. of the SPIE Symp. on Remote Sensing Conference on Image and Signal Processing for Remote Sensing V, volume 3871, pages 150–158, 1999.

[27] S.-S. Chiang, Chein-I Chang, and I. W. Ginsberg. *Unsupervised hyperspectral image analysis using independent component analysis.* In Proc. of the IEEE Int. Geosci. and Remote Sensing Symp., 2000.