

Center for Quality and Applied Statistics  
Kate Gleason College of Engineering  
Rochester Institute of Technology

**Analytical Comparison of Subpixel Target  
Detectors in Structured Models for  
Hyperspectral Images**

Peter Bajorski  
Center for Quality and Applied Statistics  
Rochester Institute of Technology  
Peter.Bajorski@rit.edu

Technical Report 2005–2

March 2005





# Analytical Comparison of Subpixel Target Detectors in Structured Models for Hyperspectral Images

Peter Bajorski

Graduate Statistics Department, Rochester Institute of Technology  
98 Lomb Memorial Drive, Rochester, NY 14623-5604  
Email: Peter.Bajorski@rit.edu

## ABSTRACT

In the current target detection literature, there are two major approaches to the evaluation of detectors' performance. One is based on theoretical calculations assuming some simple statistical models, and the other approach uses real or simulated spectral images. The former approach is too simplistic, at this point, to address practical needs. On the other hand, the latter approach does not give us a good understanding of why certain detectors work better than others in the context of specific targets and spectral images. Our goal is to initiate research that will combine these two separate approaches. In this paper, we start with a comparison of two well-known detectors—the matched filter detector (MFD) and the orthogonal subspace projection (OSP) detector. We show a surprising result that MFD always outperforms OSP in a traditional theoretical formulation of the detection problem. We also show that this theoretical formulation is not realistic in practical target detection in real spectral images. However, the obtained results suggest more realistic approaches for providing theoretical background for practical target detection.

In this paper, we also point out many detectors introduced in the literature that are equivalent to MFD or OSP detectors.

Keywords: target detection, structured model, hyperspectral image, matched filter, OSP, detection power.

## 1. INTRODUCTION

In the current target detection literature, there are two major approaches to the evaluation of detectors' performance. One is based on theoretical calculations assuming some simple statistical models, and the other approach uses real or simulated spectral images. The former approach is too simplistic, at this point, to address more practical needs. On the other hand, the latter approach does not give us a good understanding of why certain detectors work better than others in the context of specific targets and spectral images. The purpose of this paper is to initiate research that would combine these two separate approaches together. As this is not an easy task, we start with a relatively straightforward (but innovative) comparison of the two well-known detectors—the matched filter detector (MFD) and the orthogonal subspace projection (OSP) detector.

A traditional theoretical approach to target detection is to assume a certain probabilistic model and then use the appropriate detectors (test statistics) that have good properties under the assumed model. For example, the OSP detector (equivalent to the least squares estimator) seems to be a reasonable detector for the structured model (containing the target and background signatures) considered in this paper. On the other hand, the MFD is intended for a different model since it is equivalent to the least squares estimator in the structured model containing only the target signature. As such, it may seem that the MFD (using less information) would be inferior to OSP in most cases. In this paper, we investigate properties of both detectors under the same conditions, that is, for the more general structured model. We show a surprising result that the MFD always outperforms OSP in a simple detection problem often used in the current literature. We then explain why this result may not extend to practical target detection in real spectral images. These results should be treated as a first step to a theoretical analysis of more realistic detection problems.

In Section 2, we formulate the model assumed in this paper. In Section 3, we introduce the MFD and OSP detectors and also list many other detectors introduced in the literature that are equivalent to one of these two detectors. Most of the equivalencies of the detectors are already known in literature, but they are not always well known among practitioners.

We also show one new equivalency result by proving explicitly that the detector called signature space orthogonal projection classifier (SSC) in <sup>1</sup> is equivalent to the OSP detector.

The main results concerning the detection power of the MFD and OSP are shown in Section 4.

## 2. MODEL FORMULATION

Consider a set of image spectra  $\mathbf{r}_i, i=1, \dots, n$ , where  $\mathbf{r}_i$ 's are p-dimensional vectors (of reflectance or radiance, for example). For the purpose of target detection, the following structured model is often used in literature<sup>1,2,3,4,5</sup>:

$$\mathbf{r}_i = \mathbf{d} \cdot \theta_i + \mathbf{U} \cdot \boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i \quad (1)$$

where  $\mathbf{d}$  is a fixed known vector of the target spectrum,  $\mathbf{U}$  is a fixed known matrix of background spectra as columns, and  $\theta_i$ ,  $\boldsymbol{\gamma}_i$  are unknown constant and vector of abundances, respectively. The error term  $\boldsymbol{\varepsilon}_i$  is assumed to follow the multivariate normal (Gaussian) distribution  $N(\mathbf{0}, \sigma^2 \mathbf{I})$ . The normality assumption can be relaxed for some of the results in this paper, but we use it uniformly throughout the entire paper for simplicity of the presentation. The term  $\boldsymbol{\varepsilon}_i$  represents the joint effect of the noise in data and possible error in the modeling. Following a tradition in detection literature, we will sometimes refer to  $\boldsymbol{\varepsilon}_i$  as noise. The background spectra  $\mathbf{U}$  are sometimes called interference or undesired target spectra. In all subsequent notation, we suppress the index  $i$  because all methods discussed in this paper deal with one spectrum at a time.

It is a well established fact that the assumption  $Var(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}$  can be generalized to the case when  $Var(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{A}$ , where  $\mathbf{A}$  is a known matrix. The spectrum  $\mathbf{r}$  can be transformed to  $\mathbf{r}' = \mathbf{A}^{-1/2} \mathbf{r}$ , so that for the resulting  $\boldsymbol{\varepsilon}' = \mathbf{A}^{-1/2} \boldsymbol{\varepsilon}$ , the covariance matrix  $Var(\boldsymbol{\varepsilon}') = \sigma^2 \mathbf{I}$ .

The target detection problem can be defined as the hypothesis-testing problem for testing the null hypothesis  $H_0 : \theta = 0$  versus the alternative hypothesis  $H_1 : \theta > 0$ . The abundance  $\theta$  is assumed to be nonnegative according to its physical interpretation. These two hypotheses can be reformulated as follows into an equivalent form (with  $\theta > 0$ ):

$$\begin{aligned} H_0 : \mathbf{r} &= \mathbf{U} \cdot \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \\ H_1 : \mathbf{r} &= \mathbf{d} \cdot \theta + \mathbf{U} \cdot \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \end{aligned} \quad (2)$$

## 3. DETECTORS

In this section, we discuss two detectors—the matched filter detector (MFD) and the orthogonal subspace projection (OSP) detector. We also point out that many other detectors are equivalent to these two detectors. In order to simplify the presentation, we will associate the word “detector” with one of the following (where the meaning will be clear from the context):

1. An operator  $D : R^p \rightarrow R$ .
2. A vector  $\mathbf{w}$  defining a linear operator  $D : \mathbf{r} \mapsto \mathbf{w}^T \mathbf{r}$ .
3. The value of the operator  $D(\mathbf{r})$ .
4. A decision procedure indicating presence of the target when  $D(\mathbf{r}) > c$ , and its absence otherwise.

Sometimes, we will also use the statistical terminology, where the meaning 3 above is equivalent to a statistic (or an estimator if a parameter is estimated), and the meaning 4 is equivalent to a statistical test.

We would like to stress that two detectors differing by a constant (both in an additive and multiplicative sense) are equivalent. Let us assume that a detector  $D(\mathbf{r})$  indicates the presence of the target when  $D(\mathbf{r}) > c$ , where  $c$  is an appropriately chosen constant (for example, to achieve a certain false alarm probability). We can define another detector  $D'(\mathbf{r}) = a \cdot D(\mathbf{r}) + b$  indicating the presence of the target when  $D'(\mathbf{r}) > c'$ . Clearly,  $c'$  can be chosen as  $a \cdot c + b$ , so that both detectors have the same detection effect and the same properties. These statements form a sufficiently formal proof of equivalence, but a more elaborate proof is included in <sup>4</sup> (Section IV).

### 3.1. Matched filter detector (MFD)

The matched filter detector (MFD) operating on the image spectrum  $\mathbf{r}$  is defined as

$$D_{MFD}(\mathbf{r}) = \frac{\mathbf{d}^T \mathbf{r}}{\mathbf{d}^T \mathbf{d}} \quad (3)$$

The scaling of MFD is chosen so that  $D_{MFD}(\mathbf{r})$  is an unbiased estimator of  $\theta$  in the following reduced (relative to (1)) model

$$\mathbf{r} = \mathbf{d} \cdot \theta + \boldsymbol{\varepsilon} \quad (4)$$

Clearly,  $D_{MFD}(\mathbf{r})$  is the same as the least squares (and maximum likelihood) estimator of  $\theta$  in model (4).

In <sup>1</sup>, a detector called a target signature space-orthogonal projection classifier (TSC) is introduced as

$$D_{TSC}(\mathbf{r}) = \mathbf{d}^T P_{\mathbf{U}}^{\perp} P_{\mathbf{d}} \mathbf{r} \quad (5)$$

where  $P_{\mathbf{X}}$  (for  $\mathbf{X} = \mathbf{U}$  or  $\mathbf{d}$ ) is a projection operator defined by  $P_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$  and  $P_{\mathbf{X}}^{\perp} = \mathbf{I} - P_{\mathbf{X}}$  is a projection on the subspace orthogonal to  $\mathbf{X}$ . As shown in <sup>4</sup>,

$$D_{TSC}(\mathbf{r}) = \mathbf{d}^T P_{\mathbf{U}}^{\perp} \mathbf{d} (\mathbf{d}^T \mathbf{d})^{-1} \mathbf{d}^T \mathbf{r} = \mathbf{d}^T P_{\mathbf{U}}^{\perp} \mathbf{d} \cdot D_{MFD}(\mathbf{r})$$

which makes TSC equivalent to MFD.

### 3.2. Orthogonal subspace projection (OSP) detector

The orthogonal subspace projection (OSP) detector operating on the image spectrum  $\mathbf{r}$  is defined as

$$D_{OSP}(\mathbf{r}) = \frac{\mathbf{d}^T P_{\mathbf{U}}^{\perp} \mathbf{r}}{\mathbf{d}^T P_{\mathbf{U}}^{\perp} \mathbf{d}} \quad (6)$$

The OSP detector was introduced in <sup>7</sup> (with a different scaling). The connection with the regular unmixing through the least squares was pointed out in <sup>8</sup>, where the author writes that “it is not clear what advantages the use of OSP has over the direct application of spectral unmixing” (p. 1046). We agree that the two approaches are formally equivalent, however, there are some advantages of using the OSP notation (6). The reason for using the OSP name and notation here is twofold—the first one is that the OSP name is now broadly used in the detection literature, and the second is that OSP in (6) gives an explicit formula for the least squares estimator of  $\theta$ , which is convenient for some analytical considerations.

The OSP is also a special case of a matched subspace filter<sup>3</sup>. It is known that the least squares estimates are obtained through an orthogonal projection onto the subspace generated by the deterministic part of the model (1), that is,

$$P_{\mathbf{M}}\mathbf{r} = \mathbf{d} \cdot \hat{\theta}_{LS} + \mathbf{U} \cdot \hat{\gamma}_{LS} \quad (7)$$

where  $\mathbf{M} = [\mathbf{d} \ \mathbf{U}]$  is the composite matrix consisting of  $\mathbf{d}$  and all columns of  $\mathbf{U}$ . The least squares estimator  $\hat{\theta}_{LS}$  of  $\theta$  in the model (1) can also be associated with the oblique projection (see appendix), that is:<sup>9</sup>

$$\mathbf{d} \cdot \hat{\theta}_{LS} = E_{\mathbf{d}\mathbf{U}}\mathbf{r}$$

which leads to

$$\mathbf{d}^T \mathbf{d} \cdot \hat{\theta}_{LS} = \mathbf{d}^T E_{\mathbf{d}\mathbf{U}}\mathbf{r} \quad (8)$$

and together with formula (17) in the appendix, this gives

$$\hat{\theta}_{LS} = (\mathbf{d}^T \mathbf{d})^{-1} \mathbf{d}^T E_{\mathbf{d}\mathbf{U}}\mathbf{r} = (\mathbf{d}^T \mathbf{d})^{-1} \mathbf{d}^T \mathbf{d} (\mathbf{d}^\perp P_{\mathbf{U}}^\perp \mathbf{d})^{-1} \mathbf{d}^\perp P_{\mathbf{U}}^\perp \mathbf{r} = \frac{\mathbf{d}^T P_{\mathbf{U}}^\perp \mathbf{r}}{\mathbf{d}^T P_{\mathbf{U}}^\perp \mathbf{d}}$$

This is yet another confirmation that the OSP detector is equivalent to the least squares estimator of  $\theta$  in the model (1). In <sup>1</sup>, a detector called oblique subspace projection classifier (OBC) is defined as

$$D_{OBC}(\mathbf{r}) = \mathbf{d}^T E_{\mathbf{d}\mathbf{U}}\mathbf{r}$$

From (8), OBC is proportional to  $\hat{\theta}_{LS}$ , which is the same as the OSP detector. In <sup>6</sup>(Section 8.2.3), another version of the oblique projection detector is defined with additional scaling, which again is equivalent to OSP.

In <sup>1</sup>, a detector called a signature space orthogonal projection classifier (SSC) is defined as

$$D_{SSC}(\mathbf{r}) = \mathbf{d}^T P_{\mathbf{U}}^\perp P_{\mathbf{M}}\mathbf{r} \quad (9)$$

From (7),

$$D_{SSC}(\mathbf{r}) = \mathbf{d}^T P_{\mathbf{U}}^\perp (\mathbf{d} \cdot \hat{\theta}_{LS} + \mathbf{U} \cdot \hat{\gamma}_{LS}) = \mathbf{d}^T P_{\mathbf{U}}^\perp \mathbf{d} \cdot \hat{\theta}_{LS},$$

which means that the SSC detector is another scaled version of the least squares estimator and the OSP detector. This also shows that

$$D_{SSC}(\mathbf{r}) = \mathbf{d}^T P_{\mathbf{U}}^\perp \mathbf{r}$$

which is the same as the originally defined<sup>7</sup> OSP detector (also called an *a priori* OSP detector by some sources<sup>1,6</sup>). In <sup>5</sup>, a detector called the least squares orthogonal subspace projection (LSOSP) is defined using exactly the same definition as SSC in (9) (see formula (30) in <sup>5</sup>). This means that LSOSP is equivalent to OSP.

Table 1. The equivalent forms of the MFD and OSP detectors.

A Detector	Is equivalent to	Other equivalent forms
MFD	the least squares (and maximum likelihood) estimator of $\theta$ in model (4))	TSC <sup>1</sup>
OSP	the least squares (and maximum likelihood) estimator of $\theta$ in model (1))	LSOSP, <sup>5</sup> SSC, <sup>1</sup> OBC, <sup>1,9</sup> a special case of a matched subspace filter <sup>3</sup>

## 4. DETECTION POWER

Let us now discuss the relationship between hypothesis testing and practical target detection. In hypothesis testing, we need to know the distribution of  $D(\mathbf{r})$  (under the null hypothesis) in order to find the constant  $c$  that would guarantee a certain false alarm probability (probability of type I error). In practical target detection in spectral images, such distributions are usually not known, and at this point, there is no reliable method to find the appropriate constant  $c$ . So, practical target detection is based on calculating the values of  $D(\mathbf{r})$  for all pixel spectra  $\mathbf{r}$ , and associating a certain number of highest values of  $D(\mathbf{r})$  with detected targets. This is equivalent to considering certain values for  $c$  and then observing false alarm rates (as estimates of the false alarm probability) rather than calculating the desired  $c$ .

In the current target detection literature, we can identify three primary methods by which the detectors are evaluated:

1. A stochastic model for the spectrum  $\mathbf{r}$  is assumed (similar to our model (1)), and the distribution of  $D(\mathbf{r})$  is calculated, which leads to theoretical ROC curves (plots of probability of detection versus false alarm probabilities) and power curves (plots of detection power (probability of detection) versus a parameter measuring the departure from the null hypothesis).
2. An artificial set of pixel spectra is created with predetermined knowledge of target pixels. In this case, observed ROC curves (plots of detection rates versus false alarm rates) and observed power curves can be plotted. By the “rate,” we mean a relative frequency of an event (detection or false alarm) as opposed to its theoretical probability. When the set of pixel spectra is relatively small (for example, several hundred spectra as in <sup>1</sup> and <sup>5</sup>), the results resemble those for the real images, that is, statements are made about the number of pixels identified as containing the target, or simply detector values are reported. When the set of pixel spectra is very large (for example, several thousand spectra as in <sup>10</sup>) and appropriately constructed, the results can be regarded as simulated approximations of the theoretical values (such as those calculated in the first type of study).
3. A real spectral image is used, where the location of the target pixels is known. In this case, we also use observed ROC curves, while the observed power curves are not usually practical because it is unlikely to find a sufficient number of pixels with similar departures from the null hypothesis. For the same reason, the observed ROC curves are not as “clean” as those constructed in the first type of study because the target pixels are usually not homogeneous enough, that is, they may contain varying amounts of background materials.

In the study type 3, the constant  $c$  is not calculated, but instead, by changing its value, we observe a range of false alarm rates and resulting detection rates. The study type 1 resembles the statistical hypothesis testing situation, where the constant  $c$  is calculated, while the study type 3 is somewhat different because we do not necessarily need to calculate  $c$ . This point is further explored in the next section.

### 4.1. MFD versus OSP

We now want to compare the two detectors, MFD and OSP, under the assumption of model (1). Since the OSP detector is the least squares estimator of  $\theta$  in model (1), and the MFD detector does not use all information contained in model (1), it may seem that OSP should outperform MFD. We show a surprising result that the opposite is true, that is, MFD always outperforms OSP in detection problem (2). Then, we explain why the OSP may still be better in practical target detection in real images.

It is a well known result in linear models and target detection literature, that under the assumptions of model (1), the OSP statistic (6) follows the normal distribution  $N\left(\theta, \left(\mathbf{d}^T P_U^\perp \mathbf{d}\right)^{-1} \sigma^2\right)$ . The situation is a little bit more complex when model (1) is true, but we perform target detection without using the background signatures  $\mathbf{U}$ , that is, we act as if model (4) were correct. In such a case, we would use the MFD detector given by (3). According to the theory of linear models<sup>11</sup>, the MFD statistic (3) follows the normal (Gaussian) distribution

$$D_{MFD}(\mathbf{r}) \sim N\left(\theta + \mathbf{A}\boldsymbol{\gamma}, (\mathbf{d}^T \mathbf{d})^{-1} \sigma^2\right) \quad (10)$$

under the assumptions of model (1), where  $\mathbf{A} = (\mathbf{d}^T \mathbf{d})^{-1} \mathbf{d}^T \mathbf{U}$ . The same result was obtained in <sup>1</sup> (but in a more complex form) for the TSP detector (equivalent to MFD) and then was simplified in <sup>4</sup> (however, the authors could not draw our conclusions because of considering a null hypothesis different from the one used in (2)). This result means that the MFD statistic is a biased estimator of  $\theta$  under model (1). Since under the null hypothesis  $H_0$  in (2)

$$D_{MFD}(\mathbf{r}) \stackrel{H_0}{\sim} N\left(\mathbf{A}\boldsymbol{\gamma}, (\mathbf{d}^T \mathbf{d})^{-1} \sigma^2\right),$$

the decision procedure  $D_{MFD}(\mathbf{r}) > c$  achieves the probability of a false alarm  $\alpha$  when  $c = c_{MFD} = \mathbf{A}\boldsymbol{\gamma} + (\mathbf{d}^T \mathbf{d})^{-1/2} \sigma \cdot z_{1-\alpha}$ , where  $z_{1-\alpha}$  is the  $(1-\alpha)$  percentile from the standard normal (Gaussian) distribution. Since the constant vector of abundances  $\boldsymbol{\gamma}$  is unknown, we cannot find the appropriate constant  $c$  such that the decision procedure  $D_{MFD}(\mathbf{r}) > c$  would achieve a specified probability of false alarm. However, in practical target detection, we do not really calculate  $c$  but rather associate the highest values of  $D_{MFD}(\mathbf{r})$  with pixels containing the target, as explained at the beginning of Section 4. This approach is equivalent to trying different values of  $c$  and observing the properties of the decision procedure  $D_{MFD}(\mathbf{r}) > c$ . We will use the same approach here. This also explains why we can use both MFD and OSP detectors without knowing  $\sigma$ .

Using the above calculated value of  $c_{MFD}$  and property (10), we obtain the following formula for the detection power of MFD

$$P_{MFD,\alpha} = 1 - \Phi\left(z_{1-\alpha} - \frac{\theta}{\sigma} \sqrt{\mathbf{d}^T \mathbf{d}}\right) \quad (11)$$

where  $\Phi$  is the cumulative distribution function of the standard normal (Gaussian) distribution. Similar calculations for OSP lead to the following formula for its detection power

$$P_{OSP,\alpha} = 1 - \Phi\left(z_{1-\alpha} - \frac{\theta}{\sigma} \sqrt{\mathbf{d}^T P_{\mathbf{U}}^\perp \mathbf{d}}\right) \quad (12)$$

Clearly, the ROC curves for the two detectors can be found without knowing the values of abundances  $\boldsymbol{\gamma}$ . In order to simplify the analysis of detection power, we introduce the following notation:

$$B = \frac{\|\mathbf{d}\theta\|}{\sigma} = \frac{\theta}{\sigma} \sqrt{\mathbf{d}^T \mathbf{d}}$$

and  $\omega$  is the angle between  $\mathbf{d}$  and the subspace  $\langle \mathbf{U} \rangle$  generated by column vectors of  $\mathbf{U}$  ( $0^\circ \leq \omega \leq 90^\circ$ ). In practice,  $\omega \neq 0^\circ$  because  $\mathbf{d}$  is linearly independent of  $\mathbf{U}$ , and usually  $\omega \neq 90^\circ$  because two spectra (vectors of positive values) are never orthogonal except in some unrealistic cases. The angle  $\omega$  could sometimes be  $90^\circ$ , if the spectra  $\mathbf{r}$  were mean-centered, but this is not considered in this paper. The parameter  $B$  can be regarded as a signal to noise ratio (SNR) for the signal  $\mathbf{d}\theta$ . The detection power of both detectors can now be expressed as

$$P_{MFD,\alpha} = 1 - \Phi(z_{1-\alpha} - B) \quad (13)$$

$$P_{OSP,\alpha} = 1 - \Phi(z_{1-\alpha} - B \sin(\omega)) \quad (14)$$

**Theorem 1.** Under the hypotheses (2),

$$P_{OSP,\alpha} \leq P_{MFD,\alpha} \quad (15)$$

for  $\alpha \in (0,1)$ , and the equality holds if and only if  $\omega = 90^\circ$ .

Theorem 1 tells us that the detection power of MFD is always at least as high as that of OSP when using the hypotheses (2), and it is usually higher than the detection power of OSP. The only situation when the two detection powers are the same is when  $\mathbf{d}$  is orthogonal to the subspace  $\langle \mathbf{U} \rangle$ , which does not happen within the scope of this paper as explained above.

The inequality (15) tells us that the ROC curve for MFD is above the ROC curve for OSP (the curves may overlap each other in some unusual cases discussed above). Once we have the analytical formulas for the detection power, it is not very efficient to draw ROC curves because the major interest in detection is in low values of the false alarm probability. It is therefore more informative to fix a low value of false alarm probability, and draw the power curves, that is, the relationship between the detection power and a parameter that would describe the departure from the null hypothesis. Figure 1 shows such power curves (functions of the SNR parameter  $B$ ) for MFD and OSP detectors. As we know from (13), the performance of MFD is independent of the angle  $\omega$ , but based on (14), the performance of OSP is highly dependent on  $\omega$ , so we present several OSP curves for a range of  $\omega$  values.

Figure 1. Power curves for MFD and OSP detectors as functions of the SNR parameter  $B$

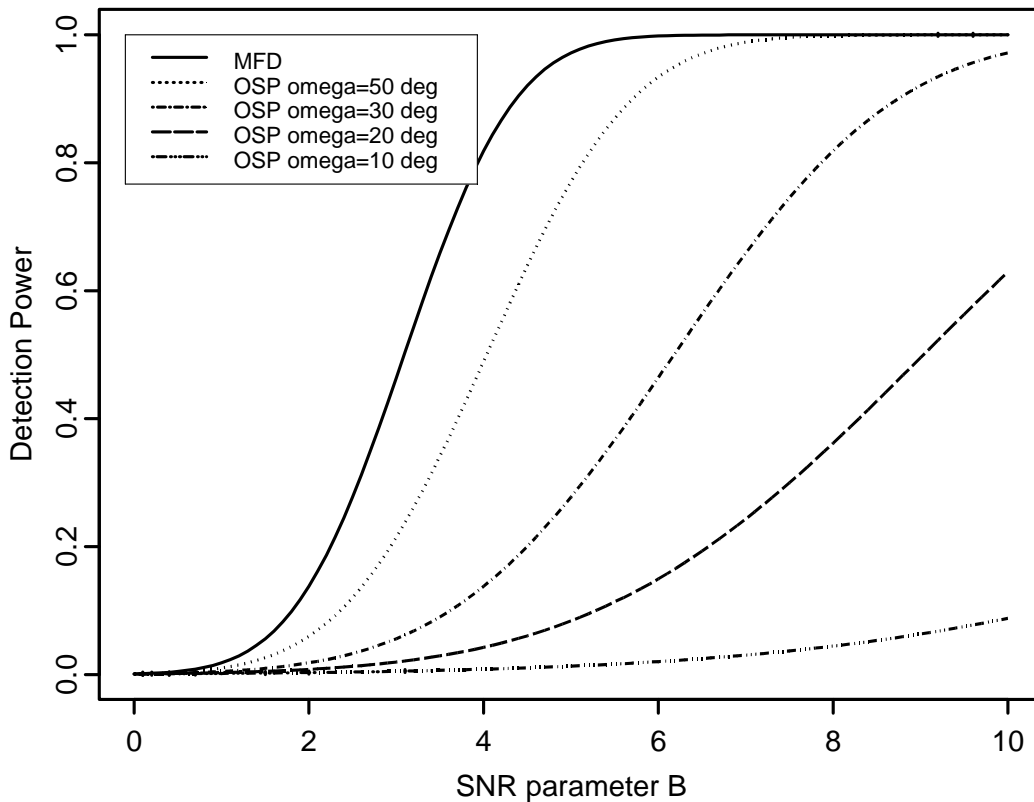


Table 2 shows some examples of possible  $\omega$  angles in practice. The five spectral signatures used in Table 2 were intensively used in papers (for example<sup>1</sup>) and in a recent book<sup>6</sup>. The spectra are shown in Figure 1.5 of the book<sup>6</sup>. Table 2 shows values of  $\omega$  when one of the signatures is used as  $\mathbf{d}$  and one of the remaining four signatures is used as the only column of  $\mathbf{U}$ . If several materials were used as columns of  $\mathbf{U}$ , then the  $\omega$  angles would be even smaller than those shown in Table 2. The maximum angle observed in Table 2 is below  $33^\circ$  and several values are below  $10^\circ$ . When using these values in the context of Figure 1, we can see that the reduction in the OSP detection power is significant.

Table 2. Angles between spectra of five materials used as an example of possible angles  $\omega$  in (14) and in Figure 1

	black-brush	creosote leaves	dry grass	red soil	sage-brush
blackbrush	0.0	10.1	14.8	23.3	3.9
creosote leaves	10.1	0.0	24.1	32.7	7.4
dry grass	14.8	24.1	0.0	12.5	17.0
red soil	23.3	32.7	12.5	0.0	25.9
sagebrush	3.9	7.4	17.0	25.9	0.0

We now show some numerical results for the three material spectra used in<sup>10</sup>. The three reflectance spectra representing barite, chalcopyrite, and pyrite were downloaded from the USGS spectroscopy lab (see the appendix Section 6.2 for details). Table 3 shows values of the  $\omega$  angle between the material spectra in a way similar to Table 2. The angle values are very small, which again indicates a low detection power of OSP compared to that of MFD.

Table 3. Angles between the three material spectra

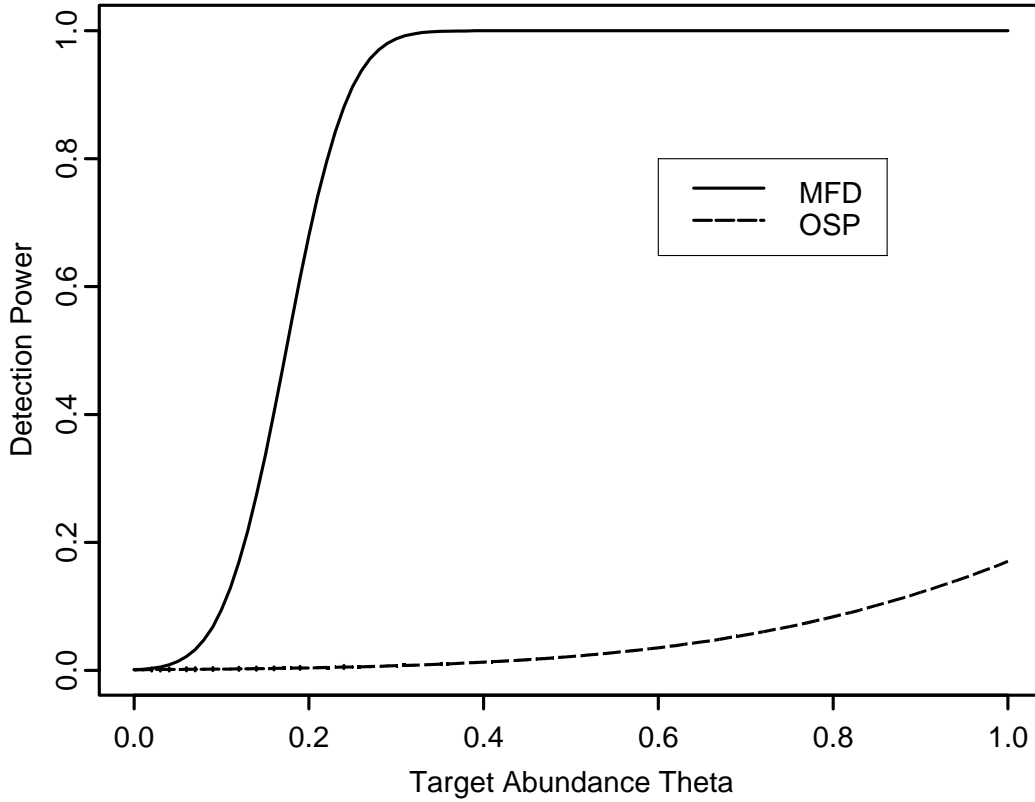
	barite	chalcopyrite	pyrite
barite	0.0	11.9	7.3
chalcopyrite	11.9	0.0	12.3
pyrite	7.3	12.3	0.0

We now want to see how the detection power of the two detectors depends on the abundance of barite as the target material, when two other materials (chalcopyrite and pyrite) are used as background materials. The angle between  $\mathbf{d}$  defined by the barite spectrum and the subspace  $\langle \mathbf{U} \rangle$  defined by chalcopyrite and pyrite is only  $6.9^\circ$ , so we can expect poor performance of OSP. Figure 2 shows the detection power of both detectors as a function of  $\theta$  assuming the false alarm probability  $\alpha = 0.001$  and  $\sigma$  such that  $\frac{\sqrt{\mathbf{d}^T \mathbf{d}}}{\sigma} = \sqrt{10^{2.5}}$  (to make it consistent with the example in<sup>10</sup>, where  $\frac{\mathbf{d}^T \mathbf{d}}{\sigma^2}$  was assumed to be 25 dB). We are trying to make this example somewhat similar to the one used in<sup>10</sup>, but there are still significant differences:

1. In<sup>10</sup>, whitened detectors CEM and TCIMF are used.
2. The target and non-target pixels generated in<sup>10</sup> are much more complicated than the two hypotheses assumed in (2).

It is relatively straightforward to use our results for the CEM and TCIMF detectors; however, the more complex hypothesis testing requires additional research.

Figure 2. The detection power of MFD and OSP detectors as a function of the abundance  $\theta$  of barite as the target material when two other materials (chalcopyrite and pyrite) are used as background materials



#### 4.2. Critique of the detection power results from a practical point of view

The previous section shows the overwhelming dominance of MFD over OSP in target detection expressed as a hypothesis testing problem (2). However, this approach is not realistic for practical target detection in real spectral images for the following reasons:

1. The value of  $\gamma$  in (2) may not be the same for  $H_0$  and  $H_1$ . For example, when the sum-to-one constraint is used,  $\gamma$  is not the same for the two hypotheses.
2. In real spectral images, we have a wide variety of non-target pixels containing different materials with varying abundance. The same consideration applies to the pixels containing the target material, where the other (background) materials may occur with varying abundance.
3. The error term properties assumed here may not hold.

We are mostly concerned about the first two reasons. We believe that by using a proper choice of background spectra in  $\mathbf{U}$  and by using whitening, it might be possible to achieve properties of the error term close to those assumed in this paper.

The results obtained in this paper allow us to understand why and in what situations the MFD and OSP detectors may or may not work well. The explanation given in <sup>10</sup> is somewhat simplistic. Even though the angle between  $\mathbf{d}$  and  $\langle \mathbf{U} \rangle$  plays an important role, the angle itself does not explain why MFD may sometimes be inferior to OSP in practical target detection. The real reason is the bias of MFD and the potentially complex relationship between the abundances of background signatures in the target-containing pixels versus the pixels not containing the target.

We treat the results presented here as a starting point to address the more complex situations as described in points 1 and 2 above.

## 5. CONCLUSIONS

We show a surprising result that MFD outperforms OSP in detection problem (2). The performance of OSP is highly dependent on the angle between the target vector and the background subspace. The smaller the angle, the lower the OSP detection power. On the other hand, MFD performance is not impacted by this angle.

We also show that the approach represented by (2) is not realistic for practical target detection in real spectral images, even though it is currently widely used in the target detection literature. However, the results shown here suggest directions for more realistic approaches for providing theoretical background for practical target detection. We plan to perform additional research to implement such approaches.

## 6. APPENDIX

### 6.1. Oblique projections

Assume that  $\mathbf{H}$  and  $\mathbf{S}$  are  $p \times m$  and  $p \times k$  full rank matrices, respectively, the composite  $p \times (m+k)$  matrix  $\mathbf{M} = [\mathbf{H} \ \mathbf{S}]$  has also full rank, and  $m+k \leq p$ . Denote by  $\langle \mathbf{H} \rangle$  a linear subspace generated by columns of  $\mathbf{H}$ . Based on our assumptions, the linear subspaces  $\langle \mathbf{H} \rangle$  and  $\langle \mathbf{S} \rangle$  are linearly independent but not necessarily orthogonal. The oblique projection is an idempotent linear operator  $E_{\mathbf{HS}}$  (that is,  $(E_{\mathbf{HS}})^2 = E_{\mathbf{HS}}$ ) whose range is  $\langle \mathbf{H} \rangle$  and whose null space is  $\langle \mathbf{S} \rangle$ . It can be shown that

$$P_{\mathbf{M}} = E_{\mathbf{HS}} + E_{\mathbf{SH}} \quad (16)$$

which is easy to see geometrically or can also be concluded from identity (20) in<sup>9</sup>. In the same paper<sup>9</sup>, it is shown that

$$E_{\mathbf{HS}} = \mathbf{H}(\mathbf{H}^\perp P_{\mathbf{S}}^\perp \mathbf{H})^{-1} \mathbf{H}^\perp P_{\mathbf{S}}^\perp \quad (17)$$

### 6.2. Details on the data downloaded from USGS spectroscopy lab

The three reflectance spectra representing barite, chalcopyrite, and pyrite were downloaded from the USGS spectroscopy lab (<http://pubs.usgs.gov/of/2003/ofr-03-395/datatable.html>). The barite file is identified as “Barite HS79.3B W1R1Bb AREF,” chalcopyrite as “Chalcopyrite S26-36 W1R1Bb AREF,” and pyrite as “Pyrite S142-1 W1R1Bc AREF.” We used only wavelengths between 0.4 and 1.8  $\mu\text{m}$  and further reduced the data set to 88 spectral bands by taking every fourth band. The three spectra are plotted in Figure 3 in the paper <sup>10</sup> (“chalpy” represents chalcopyrite, and the chalcopyrite and pyrite labels were inadvertently switched<sup>12</sup>).

### Acknowledgments

The author would like to thank Prof. Chein-I Chang for providing the spectral data used in Table 2, and Steven Johnson for providing information about the barite, chalcopyrite, and pyrite spectra used in his paper.<sup>10</sup>

## REFERENCES

1. C.-I. Chang, X.-L. 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*, vol. 36, pp. 898–912, May 1998.
2. D. Manolakis and G. Shaw. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Processing Magazine*, 19(1):29-43, January 2002.
3. L.L. Scharf, B. Friedlander, "Matched Subspace Detectors," *IEEE Trans. Signal Processing*, vol. 42, No. 8, pp. 2146–2156, August 1994.
4. S. Johnson, "The relationship between the matched filter operator and the target signature space orthogonal projection classifier," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 283–286, Jan. 2000.
5. T. M. Tu, C.-H. Chen, and C.-I Chang, "A least squares orthogonal subspace projection approach to desired signature extraction and detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 127–139, Jan. 1997.
6. C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic/Plenum Publishers, 2003.
7. J. Harsanyi and C.-I Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 779–785, July 1994.
8. J. J. Settle, "On the relationship between spectral unmixing and subspace projection," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 1045–1046, July 1996.
9. R. T. Behrens and L. L. Scharf, "Signal processing applications of oblique projections operators," *IEEE Trans. Signal Processing*, vol. 42, pp. 1413–1423, June 1994.
10. S. Johnson, "Constrained energy minimization and the target-constrained interference-minimization filter," *Optical Engineering*, vol. 42, No. 6, pp. 1850–54, June 2003.
11. N.R. Draper and H. Smith, *Applied Regression Analysis*, John Wiley & Sons, 1998.
12. S. Johnson, Personal communication.