

Hyperspectral Target Detection: Hypothesis Testing, Signal-to-Noise Ratio, and Spectral Angle Theories

Chein-I Chang[✉], *Life Fellow, IEEE*

Abstract—Hyperspectral target detection (HTD) can be generally categorized by its targets to be detected, *a priori* targets with provided known target knowledge as *a priori* target detection and *a posteriori* targets with known target signatures (spectral shapes), but unknown abundance fractions needed to be estimated as *a posteriori* target detection. As a result, target detection can be performed in three scenarios, full pure-pixel target detection corresponding to *a priori* target detection, and subpixel and mixed-pixel target detection corresponding to *a posteriori* target detection. To develop theories for these three types of target detection, this article develops three approaches. One is to rederive hypothesis testing-based detection theory using very basic statistical detection theory. Another two are new theories, signal-to-noise ratio (SNR)-based detection theory that uses SNR as a criterion to derive optimal detectors and spectral angle (SA)-based detection theory that calculates SA to perform HTD, both of which do not require prior probability distributions as hypothesis testing does. Specifically, it will be shown that many current hypothesis testing-derived likelihood ratio test (LRT)-based detectors can find their counterparts in the SNR-derived theory and the SA-derived detection theory. Finally, to evaluate the detection performance among the detectors developed from these three approaches, several effective detection measures resulting from 3-D receiver operating characteristic (ROC) analysis are used to conduct a comprehensive study and comparative analysis.

Index Terms—Likelihood ratio test (LRT), signal-to-noise ratio (SNR), spectral angle (SA), target detection (TD).

LIST OF ACRONYMS

3-D ROC	3 dimensional receiver operating characteristic
ACE	Adaptive cosine estimate.
AMD	Adaptive matched detector.
AMF	Adaptive matched filter.
ASD	Adaptive subspace detector.
AUC	Area under an ROC curve.
BKG	Background.
BS	Background suppression.
CEM	Constrained energy minimization.
DS	Data spherling/sphered.
DS-SNR	SNR is calculated in the DS-space.
GLRT	Generalized LRT.

Manuscript received February 1, 2021; revised March 3, 2021; accepted March 21, 2021. This work was supported by the Fundamental Research Funds for Central Universities under Grant 3132019341.

The author is with the Center for Hyperspectral Imaging in Remote Sensing (CHIRS), Information and Technology College, Dalian Maritime University, Dalian 116026, China, and also with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD 21250 USA (e-mail: cchang@umbc.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TGRS.2021.3069716>.

Digital Object Identifier 10.1109/TGRS.2021.3069716

GDS-SNR	Generalized DS-SNR.
GK-SNR	Generalized K-SNR.
GR-SNR	Generalized R-SNR.
GSNR	Generalized SNR.
HTD	Hyperspectral target detection.
K/K	Covariance matrix.
K-SNR	SNR is calculated in the K-space.
K-space	Data space is whitened by $\mathbf{K}^{-1/2}$.
K-SA	calculated in the K-space.
K-SA ²	SA^2 calculated in the K-space.
LRT	Likelihood ratio test.
MD	Mahalanobis distance.
MLE	Maximum likelihood estimate.
NAMD	Normalized adaptive matched detector.
NAMF	Normalized adaptive matched filter.
NGDS-SNR	Normalized GDS-SNR.
NGK-SNR	Normalized GK-SNR.
NGR-SNR	Normalized GR-SNR.
NLRT	Normalized likelihood ration test.
NMF	Normalized matched filter.
NOSP-GLRT	Normalized OSP-GLRT.
NP	Neyman–Pearson.
OSP	Orthogonal subspace projection.
R/R	Correlation matrix.
R-space	Data space is whitened by $\mathbf{R}^{-1/2}$.
R-SNR	SNR is calculated in the R-space.
R-SA ²	SA^2 calculated in the R-space.
ROC	Receiver operating characteristic.
SA	Spectral angle.
SNR	Signal-to-noise ratio.
TD	Target detectability.

I. INTRODUCTION

Hyperspectral target detection (HTD) is a fundamental task in hyperspectral data exploitation [1]–[3]. Owing to very fine spectral resolution, a hyperspectral imaging sensor is capable of uncovering many subtle substances which cannot be identified by prior knowledge or visual inspection. For HTD, a detection problem is generally formulated by a binary hypothesis testing problem with noise assumed to be additive Gaussian. Unfortunately, a hyperspectral image generally has a very complicated BKG including many unknown and unidentified material substances and interferers other than noise. As a result, one of major challenging issues arising in hyperspectral data exploitation is how to deal with BKG effects.

In addition to the BKG issue, the statistical signal detection theory generally considers a single-target signal detection in

the sense that the signal to be detected is one signal source. However, in hyperspectral imaging, target signal sources can be categorized into three types, namely, full pure pixel targets, subpixel targets, and mixed pixel targets, each of which has its own particular characteristics and requires a separate theory to derive its detector [2, Ch. 2]. Of particular interest are subpixel targets and mixed pixel targets where their target signatures (i.e., signal shapes/envelops considered in communications/signal processing) are assumed to be known but their abundance fractions (i.e., signal intensity considered in communications/signal processing) are unknown and need to be estimated prior to target detection.

A classic approach is to formulate signal detection as a binary hypothesis testing problem where the NP detection theory is used to maximize the detection probability of a detector while imposing a constraint on its false alarm probability [4]. To extend the NP detection to HTD in the three types of target pixels, the full pure pixel targets and subpixel/mixed pixel targets, the binary hypothesis testing is formulated with a binary simple hypothesis testing problem corresponding to pure pixel target detection and a binary composite hypothesis testing problem corresponding to subpixel/mixed pixel target detection where likelihood ratio test (LRT)-based and generalized LRT (GLRT) detectors are derived, respectively. This article rederives LRT/GLRT-based detectors using the basic statistical detection theory in [4, Ch. 2]. Specifically, it separates subpixel/mixed pixel targets from BKG as individual target pixels so that target detection can be performed by detecting the target abundance fractions.

The hypothesis testing-derived LRT/GLRT detectors described above require a specific probability distribution under each hypothesis, which is generally very difficult to obtain. Under this circumstance, the Gaussian distribution is always assumed for mathematical tractability. In hyperspectral imagery, such Gaussian assumption is usually not valid due to the presence of many unknown subtle substances and interferer in the data. To address this issue, a new approach, called SNR-based detection theory, also known as deflection-based detection theory [5] is developed. Its idea is originated from the OSP developed in [6] which an OSP-based signal detector using SNR as a criterion for optimality was derived. Interestingly, such SNR-derived OSP detector was also shown to be the Gaussian maximum likelihood estimator in [7] and [8] if the probability distributions were assumed to be multivariate Gaussian. Given the fact that the LRT-based detectors are generally derived from hypothesis testing problems with Gaussian noise assumption, the work in [5] and [6] provides the base of the SNR-based detection theory derived in this article. Most importantly, the hypothesis testing-derived LRT/GLRT can be shown to have their counterparts in the SNR-detection theory. In particular, the relationship between LRT/GLRT-derived detectors and SNR/GSNR-derived detectors is further explored to show that LRT/GLRT-based directors are essentially the same as SNR-derived detectors.

The abovementioned hypothesis testing-based and SNR-based approaches give rise to linear matched filters (MFs) by using various matching signals to perform signal detection.

Along this line, a third approach is further developed as another type of MF-based approach. It is called SA-detection theory, which calculates the SA between the target signature to be matched with the data sample being processed for detection. Furthermore, to take care of the data correlation issue, three types of data whitening techniques, DS, covariance matrix-based whitening and correlation matrix-based whitening, are particularly developed as a preprocessing step before the SA calculation. As a result, SA operated in such whitened data space performs surprisingly well and generally does better than statistical model-based LRT/GLRT detectors.

To evaluate detection performance, the common 2-D ROC analysis has been widely used as an analysis tool for this purpose. Unfortunately, the traditional 2-D ROC curve has been shown to be not effective [9]. In this case, several detection measures derived from the 3-D ROC analysis in [9] are also used for detection performance analysis.

Several novelties of this article are summarized as follows.

- 1) Three scenarios of HTD, fully occupied pure-pixel target detection, subpixel, and mixed-pixel target detection, are developed for hypothesis testing theory to derive their respective detectors, specifically, mixed pixel target detection which includes OSP to derive OSP-based target detectors for mixed pixel target detection.
- 2) A new simple detection approach derived from [4] is used to rederive the HTD theory.
- 3) A new SNR-based signal detection theory is developed for HTD to show that LRT/GLRT-derived detectors are indeed SNR-derived detectors where no probability distributions are assumed. In particular, the well-known subpixel target detection technique, CEM [10], [11] can be also shown to be a special case of a detector derived from the SNR detection theory.
- 4) A new relationship between LRT/GLRT-derived and SNR/GSNR-derived detectors is explored and established where LRT/GLRT can be interpreted as special cases under the umbrella of the SNR-detection theory.
- 5) A new SA-based signal detection theory in conjunction with data whitening is developed for HTD where the adaptive cosine estimator (ACE) in [12] can be also shown to be a special case of an SA-detector.

Most importantly, given an assumption that BKG can be characterized by the second-order data statistics, removing the BKG effect can be done by decorrelating of the second-order data statistics. This is similar to taking inverse of the covariance matrix in a Gaussian distribution to decorrelate data. To accomplish this goal, three data whitening methods are proposed in this article to perform data decorrelation. All of these three methods have been used to derive LRT, SNR, and SA detectors and are described in the following.

- 1) DS removes the first two orders of data statistic, i.e., sample mean and sample variances/covariances.
- 2) Covariance matrix \mathbf{K} -data whitening uses $\mathbf{K}^{-1/2}$ to decorrelate the second-order statistics but not remove variances.
- 3) Correlation matrix \mathbf{R} -data whitening uses $\mathbf{R}^{-1/2}$ to decorrelate the second-order statistics by removing covariances but not variances.

The remainder of this article is organized as follows. Section II reviews the hypothesis testing-based detection theory in great detail. Section III describes novelties derived from this article. Section IV rederives hypothesis testing-based detection theory for three types of target pixels, namely, full pure pixels, subpixels, and mixed pixels. Section V presents a new detection theory based on SNR, called SNR-based detection theory. Section VI explores the relationship between the hypothesis testing-based detection theory and SNR-based detection theory and further shows that the hypothesis testing-derived detection theory can be interpreted as its special case. Section VII presents a simple nonlinear MF theory, called SA-based detection theory, which also interprets some well-known detectors as its special cases. Section VIII introduces several detection measures derived from 3-D ROC curves, which will be used to evaluate detection performance. Section IX conducts a very comprehensive and comparative study and analysis among detectors derived from hypothesis testing-based, SNR-based and SA-based detection theories. Finally, Section X concludes and summarizes the contributions made in this article.

II. RELATED WORKS

A general approach to signal detection in noise is to formulate a signal detection problem via a binary hypothesis testing problem where the null hypothesis, H_0 and the alternative hypothesis, H_1 represent noise and a known signal in additive noise with the two hypotheses governed by known probability distributions, $p_0(\mathbf{r})$ and $p_1(\mathbf{r})$, respectively. According to [4], the sought optimal detector is LRT, which is $p_1(\mathbf{r})/p_0(\mathbf{r})$ given by a ratio of $p_1(\mathbf{r})$ to $p_0(\mathbf{r})$. When a signal to be detected is specified by a known target signature (i.e., target spectral shape) parameterized by unknown signal intensity, referred to as target abundance fraction, the binary simple hypothesis testing problem is extended to a binary composite hypothesis testing problem where LRT is also generalized to GLRT. Accordingly, a binary simple hypothesis testing problem is formulated to correspond to pure pixel target detection, while a binary composite hypothesis testing problem is being formulated to correspond to subpixel/mixed pixel target detection. If there are multiple targets to be detected, a binary hypothesis testing problem can be extended to a multiple-hypothesis testing problem where each hypothesis represents a particular signal to be detected [4], [13].

In the past, many works on HTD have followed the hypothesis testing approach to design many detectors for hyperspectral images [14]–[24]. For example, Robey *et al.* [14] and Scharf *et al.* [15] developed a GLRT as AMF detector as well as subspace-based GLRT for the target present in Gaussian noise with unknown covariances, respectively. Kraut and Scharf [16] also discussed the relationships among matched subspace detector (MSD), AMF, and GLRT. Kraut *et al.* [17] extended the theory of GLRT to adapt the MSDs to unknown noise covariance matrices, referred to as ASDs. Manolakis *et al.* [18] conducted a comparative analysis of hyperspectral AMF detectors with ACE and ASD derived from unstructured BKG and the structured BKG,

respectively. Manolakis *et al.* [19] further used a linear mixing model to perform subpixel/mixed-pixel detection. Manolakis and Shaw [20] provided a tutorial review on HTD. Specifically, Manolakis *et al.* [21] discussed spectral variability models and subpixel target detection in unstructured and structured BKG. Manolakis *et al.* [22] also compared performance among AMF, ACE, ASD, and OSP. Manolakis *et al.* [12] particularly investigated the subpixel detection problem by a model in an unstructured BKG

$$\mathbf{r} = \alpha \mathbf{t} + (1 - \alpha) \mathbf{b} \text{ in an unstructured BKG} \quad (1a)$$

known as the replacement signal model or target-BKG mixture model and a model in a structured BKG

$$\mathbf{r} = \alpha \begin{bmatrix} \mathbf{t} \\ \mathbf{b} \end{bmatrix} + \mathbf{n} \text{ in a structured BKG} \quad (1b)$$

known as subspace model, both of which can be used to detect the subpixel target \mathbf{t} in terms of its unknown abundance fraction α present in the BKG-corrupted data sample \mathbf{r} . Manolakis *et al.* [23] and Golikov *et al.* [24] further explored the subpixel target detection in unstructured and structured BKG as discussed in [12] and [20]–[22]. A major issue arising in (1a) and (1b) is how to deal with the unknown abundance fraction α which is needed to be estimated.

In many practical applications, *a priori* information about the targets of interest is usually not available or only partially available such as subpixel or mixed pixel targets. A common approach is to utilize statistical models to characterize BKG so that a hypothesis testing-based signal detection model can be still applicable. In this case, a binary simple hypothesis testing is extended to a binary composite hypothesis testing formulation where the alternative hypothesis, H_1 specifies a known signal spectrum (spectral shape) with unknown signal intensity (signal abundance fraction), while the null hypothesis, H_0 represents BKG in which the target is embedded. With this formulation, a multivariate Gaussian distribution model proposed in [14] and [15] allows maximum likelihood estimation (MLE) to be carried out to estimate α . This approach leads to ACE from (1a) and ASD from (1b) along with their variants [12], [14]–[23].

III. NOVELTIES

This article looks into a quite different approach from the above statistical BKG model-based approach to deal with a subpixel/mixed target detection problem. Instead of using signal models specified by (1a) and (1b), we simply consider a subpixel target as $\alpha \mathbf{t}$ embedded in an additive BKG, denoted by \mathbf{b} which also includes noise. In particular, this simple model can be further extended to the mixed pixel target detection with $\alpha \mathbf{t}$ replaced by mixed signal $\mathbf{t} = \mathbf{M}\alpha$ where \mathbf{M} is a target signature matrix that forms a mixed signal \mathbf{t} . To unmix the signal $\mathbf{t} = \mathbf{M}\alpha$, OSP developed in [6] can be used for this purpose. Accordingly, the GLRT-derived subpixel detectors can be further extended to OSP-GLRT detectors for mixed-pixel target detection which has not been investigated in the past.

Over the past years, the hypothesis testing-derived LRT approach has dominated HTD. One major issue encountered in this approach is its Gaussian assumption made on noise

or BKG in [12] and [14]–[24]. In real hyperspectral imagery, such assumption is generally not valid. To reinvent the wheel, this article proposes a new SNR-based detection theory for HTD which does not require prior knowledge of probability distributions or hypothesis testing problems. It simply assumes signals to be detected embedded in BKG additively, that is signals and BKG are uncorrelated. According to [4], [5], LRT-derived detectors can be also characterized by SNR. Specifically, the multivariate Gaussian noise-based MLE used by GLRT-derived detectors can now be replaced by MD to derive SNR-GLRT detectors without making Gaussian assumption. Besides, to achieve what the Gaussian covariance matrix used in LRT does, data whitening is included as a preprocessing step before calculation of SNR. Three data whitening methods are particularly designed, DS, sample covariance matrix \mathbf{K} -data whitening, and sample correlation matrix \mathbf{R} -data whitening. Surprisingly, all the hypothesis testing-derived LRT/GLRT detectors can be also derived under the context of SNR/GSNR using data whitening techniques. It is also interesting to show that the well-known subpixel target detection technique, CEM developed in [10] and [11] is actually one of SNR-derived detectors.

In addition to the hypothesis testing-based and SNR-based detection theories, a third new approach, to be called SA-based detection theory, is also derived in this article. In analogy with the SNR-based detection, three data whitening methods are also implemented prior to calculation of SA. Coincidentally, ACE derived from the signal model (1a) in the unstructured BKG can be also shown to be one of SA-detectors operating in the covariance matrix-whitened data.

Although there are also many efforts in improving HTD techniques such as hierarchical CEM in [25], ensemble CEM [26], kernel-based SA [27], hypothesis testing-based sparse representation [28] and hypothesis testing-based low-rank-sparse matrix [29], extending these works by the proposed work is feasible. Unfortunately, their discussions go way beyond the scope of this article and are not discussed here.

Finally, a common practice to evaluate the effectiveness of a detector is to use the area under a 2-D ROC curve of (P_D , P_F), denoted by $AUC_{(D,F)}$, for performance analysis. It has been shown in [9] that two detectors may have very close $AUC_{(D,F)}$ values, but they do have a significant difference in detection maps. Such phenomena were also witnessed by the experiments conducted in this article. Therefore, it will be extremely difficult to rely on one single $AUC_{(D,F)}$ to differentiate one detector from another. In this article, the detection measures derived in [9] will be used for the performance evaluation of all test detectors. It turns out that no single detector can perform universally optimal on all the detection measures. Nevertheless, their performance can be evaluated according to their functionalities, target detectability (TD), BKG suppressibility (BS), signal-to-noise probability ratio (SNPR), and overall detection probability (ODP).

IV. HYPOTHESIS TESTING-BASED HTD

This section presents a hypothesis testing-based HTD theories for three types of scenarios in HTD, pure-pixel target

detection, subpixel target detection, and mixed-pixel target detection.

A. Pure Target Detection

A standard target detection problem for HTD can be formulated by a binary simple hypothesis testing problem as follows:

$$\begin{aligned} H_0 : \mathbf{r} &= \mathbf{b} \\ \text{versus} \\ H_1 : \mathbf{r} &= \mathbf{t} + \mathbf{b} \end{aligned} \quad (2)$$

where \mathbf{r} and \mathbf{t} are L -dimensional vectors with L being the number of spectral bands and \mathbf{b} is the BKG signature. In (2) the target signal \mathbf{t} is assumed to be known and deterministic. This type of target detection is referred to as *a priori* target detection.

The hypotheses H_0 and H_1 are called “null hypothesis” which represents signal absence and “alternative hypothesis” which represents signal presence, respectively, each of which is governed by their own probability density functions, $p_0(\mathbf{r})$ and $p_1(\mathbf{r})$. Assume that the prior probabilities of H_0 and H_1 are given by $P(H_0)$ and $P(H_1)$, respectively. Also, let c_{ij} be the cost incurred by a decision saying H_i when H_j is true.

The solution to (1) can be found in [4]. It is the Bayes detector $\delta^{\text{Bayes}}(\mathbf{r})$ which is a likelihood ratio test (LRT), $\Lambda(\mathbf{r})$ given by

$$\delta^{\text{LRT}}(\mathbf{r}) = \begin{cases} 1, & \text{if } \Lambda(\mathbf{r}) \geq \tau \\ 0, & \text{if } \Lambda(\mathbf{r}) < \tau \end{cases} \quad (3)$$

where $\Lambda(\mathbf{r}) = p_0(\mathbf{r})/p_1(\mathbf{r})$, the threshold τ is given by $\tau = ((c_{10} - c_{00})P(H_0)/(c_{01} - c_{11})P(H_1))$. Using (3) P_D and P_F can be calculated by

$$\begin{aligned} P_D &= \int_{\Lambda(\mathbf{r}) \geq \tau} p_1(\mathbf{r}) d\mathbf{r} \\ P_F &= \int_{\Lambda(\mathbf{r}) \geq \tau} p_0(\mathbf{r}) d\mathbf{r}. \end{aligned} \quad (4)$$

If $p_0(\mathbf{r})$ of \mathbf{b} in (2) is assumed to be Gaussian distributed with the mean $\mu_{\mathbf{b}}$ and covariance matrix \mathbf{K} , then (3) can be derived by

$$\begin{aligned} \log \Lambda(\mathbf{r}) &= (1/2) [(\mathbf{r} - \mu_{\mathbf{b}})^T \mathbf{K}^{-1} (\mathbf{r} - \mu_{\mathbf{b}}) - (\mathbf{r} - \mu_{\mathbf{b}} - \mathbf{t})^T \mathbf{K}^{-1} (\mathbf{r} - \mu_{\mathbf{b}} - \mathbf{t})] \\ &= \mathbf{t}^T \mathbf{K}^{-1} (\mathbf{r} - \mu_{\mathbf{b}}) - (1/2) \mathbf{t}^T \mathbf{K}^{-1} \mathbf{t} \geq \log \tau \\ &= \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r} \geq \mathbf{t}^T \mathbf{K}^{-1} \mu_{\mathbf{b}} + (1/2) \mathbf{t}^T \mathbf{K}^{-1} \mathbf{t} + \log \tau = \tau'. \end{aligned} \quad (5)$$

The detector in (5) is a linear MF that uses the known target signal \mathbf{t} as the matching signal. In this case, (3) becomes

$$\delta^{\text{LRT}}(\mathbf{r}) = \begin{cases} 1, & \text{if } \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r} \geq \tau' \\ 0, & \text{if } \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r} \geq \tau'. \end{cases} \quad (6)$$

According to [4, pp. 56–57], the SNR for (6) is obtained by

$$\max_{\mathbf{r}} \text{SNR}^{\text{LRT}}(\mathbf{r}) = \text{var}[\delta^{\text{LRT}}(\mathbf{r}) | H_j] = \mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}. \quad (7)$$

Since the τ in (3) and τ' in (6) are determined by the prior cost c_{ij} and the prior knowledge of probabilities, $P(H_0)$ and

$P(H_1)$, $p_0(\mathbf{r})$ and $p_1(\mathbf{r})$, they are not generally available in practical applications. To resolve this issue, the NP detection theory is widely used to replace τ and τ' with a constraint β imposed on P_F in (4) and then maximize the detection probability P_D in (4) as follows:

$$\max_{\delta} \{P_D(\delta(\mathbf{r}))\} \text{ subject to } P_F(\delta(\mathbf{r})) \leq \beta. \quad (8)$$

The resulting detector is called NP detector which is shown in [4] to be the same LRT given by

$$\delta^{\text{NP}}(\mathbf{r}) = \begin{cases} 1, & \text{if } \Lambda(\mathbf{r}) \geq \tau \\ 0, & \text{if } \Lambda(\mathbf{r}) < \tau \end{cases} \quad (9)$$

where the threshold τ can be found by solving

$$P_F(\delta(\mathbf{r})) = \int_{\Lambda(\delta(\mathbf{r})) \geq \tau} p_0(\mathbf{r}) d\mathbf{r} = \beta \Leftrightarrow \tau = P_F^{-1}(\beta). \quad (10)$$

The major advantage of using the NP detector in (9) over (3) or (6) is no prior knowledge required. Instead, the NP detector is determined by its level of significance β which is specified by real-world applications.

B. Subpixel Target Detection

A subpixel target detection problem can be formulated by a binary composite hypothesis testing problem as follows:

$$\begin{aligned} H_0 : \mathbf{r} &= \mathbf{b} \\ \text{versus} \\ H_1 : \mathbf{r} &= \alpha \mathbf{t} + \mathbf{b} \end{aligned} \quad (11)$$

where \mathbf{r} is an L -dimensional vector with L being the number of spectral bands and \mathbf{b} is the BKG signature vector with its probability distribution described by a Gaussian distribution with the mean μ_b and covariance matrix \mathbf{K} . Specially, in (11), $\alpha \mathbf{t}$ is a subpixel target with \mathbf{t} being specified by *a priori* target spectral signature, which is also an L -dimensional vector and α is an unknown target intensity parameter as abundance fraction of \mathbf{t} , which is assumed to be unknown and needs to be estimated. Since the parameter α needs to be estimated before (11) is carried out for target detection, (11) is referred to as *a posteriori* target detection. It should be noted that (11) is different from the replacement signal model in (1a) in the sense that the subpixel target in (11) is not mixed by the BKG signature \mathbf{b} as it is in (1a). In this case, $p_0(\mathbf{r}) \sim N(\mu_b, \mathbf{K})$ and $p_\alpha(\mathbf{r}) \sim N(\alpha \mathbf{t} + \mu_b, \mathbf{K})$ are two probability distributions under H_0 and H_1 , respectively. The solution to (11) is GLRT given by

$$\delta^{\text{GLRT}}(\mathbf{r}) = \frac{\max_\alpha p_\alpha(\mathbf{r})}{p_0(\mathbf{r})} = \frac{p_{\hat{\alpha}_{\text{MLE}}(\mathbf{r})}(\mathbf{r})}{p_0(\mathbf{r})}. \quad (12)$$

Or alternatively

$$\begin{aligned} \Lambda(\mathbf{r}) &= \frac{\max_\alpha p_\alpha(\mathbf{r})}{p_0(\mathbf{r})} \\ &= \frac{\exp[-(\mathbf{r} - \hat{\alpha}_{\text{MLE}}(\mathbf{r})\mathbf{t} - \mu_b)^T \mathbf{K}^{-1}(\mathbf{r} - \hat{\alpha}_{\text{MLE}}(\mathbf{r})\mathbf{t} - \mu_b)]}{\exp[-(\mathbf{r} - \mu_b)^T \mathbf{K}^{-1}(\mathbf{r} - \mu_b)]} \\ &\Leftrightarrow \log \Lambda(\mathbf{r}) \\ &= 2\hat{\alpha}_{\text{MLE}}(\mathbf{r})\mathbf{t}\mathbf{K}^{-1}(\mathbf{r} - \mu_b) \end{aligned} \quad (13)$$

where

$$\begin{aligned} \hat{\alpha}_{\text{MLE}}(\mathbf{r}) &= \arg \left\{ \max_\alpha p_\alpha(\mathbf{r}) \right\} \\ &= \arg \left\{ \min_\alpha (\mathbf{r} - \alpha \mathbf{t} - \mu_b)^T \mathbf{K}^{-1}(\mathbf{r} - \alpha \mathbf{t} - \mu_b) \right\}. \end{aligned} \quad (14)$$

To solve (14)

$$\begin{aligned} \frac{\partial \left[(\mathbf{r} - \alpha \mathbf{t} - \mu_b)^T \mathbf{K}^{-1}(\mathbf{r} - \alpha \mathbf{t} - \mu_b) \right]}{\partial \alpha} \Big|_{\hat{\alpha}_{\text{MLE}}(\mathbf{r})} &= 0 \\ \Rightarrow \hat{\alpha}_{\text{MLE}}(\mathbf{r}) &= \frac{\mathbf{t}^T \mathbf{K}^{-1}(\mathbf{r} - \mu_b)}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}}. \end{aligned} \quad (15)$$

Substituting (15) in (6) yields

$$\delta^{\text{GLRT}}(\mathbf{r}) = \begin{cases} 1; & \text{if } \hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t} \mathbf{K}^{-1}(\mathbf{r} - \mu_b) \geq \tau' \\ 0; & \text{if } \hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t} \mathbf{K}^{-1}(\mathbf{r} - \mu_b) < \tau' \end{cases}$$

and

$$[\hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t}]^T \mathbf{K}^{-1}(\mathbf{r} - \mu_b) = \frac{1}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}} (\mathbf{t}^T \mathbf{K}^{-1}(\mathbf{r} - \mu_b))^2 \quad (16)$$

with $\tau' = (1/2) \log \tau$.

The GLRT in (16) is also called AMF in [14]. The square root of (16) is called MF [23]. Specifically, GLRT can be also shown in [23] to be an anomaly detector, referred to as Reed–Xiaoli’s anomaly detector (RXD) [24].

In analogy with (7), the SNR of (16) can be also defined as

$$\begin{aligned} \text{SNR}(\delta^{\text{GLRT}}(\mathbf{r})) &= \text{var}[\delta^{\text{GLRT}}(\mathbf{r}) | H_j] \\ &= E \left[\left((\hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t} + \mu_b) - \mu_b \right)^T \mathbf{K}^{-1}(\mathbf{r} - \mu_b) \right]^2 | H_j \\ &= E \left[\left([\hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t}]^T \mathbf{K}^{-1}(\mathbf{r} - \mu_b) \right)^2 | H_j \right] \\ &= (\hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t})^T \left[E \left(\mathbf{K}^{-1}(\mathbf{r} - \mu_b) \right) \left(\mathbf{K}^{-1}(\mathbf{r} - \mu_b) \right)^T \right] (\hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t}) \\ &= (\hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t})^T \mathbf{K}^{-1} E[(\mathbf{r} - \mu_b)(\mathbf{r} - \mu_b)^T] \mathbf{K}^{-1} (\hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t}) \\ &= (\hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t})^T \mathbf{K}^{-1} (\hat{\alpha}_{\text{MLE}}(\mathbf{r}) \mathbf{t}) = \frac{(\mathbf{t}^T \mathbf{K}^{-1}(\mathbf{r} - \mu_b))^2}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}} \\ &\text{since } E[(\mathbf{r} - \mu_b)(\mathbf{r} - \mu_b)^T] = \mathbf{K} \end{aligned} \quad (17)$$

which is identical to (16) with $\hat{\alpha}_{\text{MLE}}(\mathbf{r})$ is given by (15). Unlike SNR in (7) which is a constant, The SNR in (17) is not but rather a function of the data sample \mathbf{r} . This is because the signal abundance α in (11) is unknown and needs to be estimated from \mathbf{r} .

C. Mixed Target Detection

A mixed-pixel target detection problem can be formulated by a binary composite hypothesis testing problem as follows:

$$\begin{aligned} H_0 : \mathbf{r} &= \mathbf{b} \\ \text{versus} \\ H_1 : \mathbf{r} &= \mathbf{M}\alpha + \mathbf{b} \end{aligned} \quad (18)$$

where \mathbf{r} is an L -dimensional vector with L being the number of spectral bands, $\mathbf{M} = [\mathbf{t}_1 \mathbf{t}_2, \dots, \mathbf{t}_p]$ is a target signature matrix consisting of p target signatures, $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_p$ along with $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p)^T$ being an unknown p -dimensional abundance fractional vector of \mathbf{M} and \mathbf{b} is the BKG signature vector with its probability distribution described by a Gaussian

TABLE I
LIST OF LRT-BASED DETECTORS DERIVED FROM HYPOTHESIS TESTING THEORY

LRT-based methods	Full pure pixel detectors	Subpixel detectors	Mixed pixel detectors	Equation
LRT	$\delta^{\text{LRT}}(\mathbf{r}) = \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}$	NA	NA	(6)
GLRT=AMF	NA	$\delta^{\text{GLRT}}(\mathbf{r}) = \left(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t} \right)^{-1} \left(\mathbf{t}^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b) \right)^2$ $= \delta^{\text{AMF}}(\mathbf{r})$	NA	(16)
OSP-GLRT	NA	NA	$\delta^{\text{OSP-GLRT}}(\mathbf{r}) = \left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p \right)^{-1} \left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} P_{\mathbf{U}_p}^T (\mathbf{r} - \boldsymbol{\mu}_b) \right)^2$	(23)

distribution with μ and covariance matrix \mathbf{K} . Once again, the parameter vector α is unknown and needed to be estimated prior to (18). So, (18) is also referred to as *a posteriori* target detection. In this case, $\mathbf{M}\alpha$ is a mixed-pixel target to be detected and LRT in (6) cannot directly applicable since it does not make sense to detect a mixed-pixel target. Also comparing (18) to (1b), there is a significant difference of (18) from the subspace model in (1b) in the sense that the mixed signal in (18) is mixed by set of p target signatures, while the target signature \mathbf{t} and BKG signature \mathbf{b} in (1b) are represented by a single state vector.

To solve (18), we follow the OSP approach developed in [6] by separating a signal of interest to be detected, say \mathbf{t}_p as the desired target signature, and the remaining $(p-1)$ signatures, $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_{p-1}$ as the undesired target signatures. With this formulation, we can first annihilate the undesired target signatures $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_{p-1}$ from a data sample \mathbf{r} prior to detection of \mathbf{t}_p . In doing so, we define an orthogonal subspace projector given by $P_{\mathbf{U}_p}^\perp$

$$P_{\mathbf{U}_p}^\perp = \mathbf{I} - \mathbf{U}_p \mathbf{U}_p^\# = \mathbf{I} - \mathbf{U}_p (\mathbf{U}_p^T \mathbf{U}_p)^{-1} \mathbf{U}_p^T \quad (19)$$

where $\mathbf{U}_p = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_{p-1}]$ is an undesired target signature matrix formed by $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_{p-1}$ and $\mathbf{U}_p^\#$ is the pseudo inverse of \mathbf{U}_p given by $(\mathbf{U}_p^T \mathbf{U}_p)^{-1} \mathbf{U}_p^T$. Applying (19) to the binary composite hypothesis testing problem in (18) yields the following U-annihilated binary composite hypothesis testing problem:

$$\begin{aligned} H_0 : \tilde{\mathbf{r}} &= P_{\mathbf{U}_p}^\perp \mathbf{b} \\ \text{versus} \\ H_1 : \tilde{\mathbf{r}} &= P_{\mathbf{U}_p}^\perp (\mathbf{M}\alpha) + P_{\mathbf{U}_p}^\perp \mathbf{b}. \end{aligned} \quad (20)$$

Since $\tilde{\mathbf{t}} = P_{\mathbf{U}_p}^\perp \mathbf{t} = \mathbf{t}_p$ and $\tilde{\mathbf{b}} = P_{\mathbf{U}_p}^\perp \mathbf{b}$, (20) becomes

$$\begin{aligned} H_0 : \tilde{\mathbf{r}} &= \tilde{\mathbf{b}} \\ \text{versus} \\ H_1 : \tilde{\mathbf{r}} &= \alpha_p \mathbf{t}_p + \tilde{\mathbf{b}} \end{aligned} \quad (21)$$

which is reduced to exactly the subpixel target detection specified by (11). By replacing \mathbf{r} and \mathbf{t} in (16) and (17) with

$$\begin{aligned} \tilde{\mathbf{r}} &= P_{\mathbf{U}_p}^\perp \mathbf{r}, \tilde{\mathbf{t}} = P_{\mathbf{U}_p}^\perp \mathbf{t} = \mathbf{t}_p \\ \tilde{\mathbf{K}} &= (1/N) \sum_{n=1}^N (\tilde{\mathbf{r}}_n - \tilde{\boldsymbol{\mu}}_b)(\tilde{\mathbf{r}}_n - \tilde{\boldsymbol{\mu}}_b)^T \\ &= (1/N) \sum_{n=1}^N \left(P_{\mathbf{U}_p}^\perp \mathbf{r}_n - P_{\mathbf{U}_p}^\perp \boldsymbol{\mu}_b \right) \left(P_{\mathbf{U}_p}^\perp \mathbf{r}_n - P_{\mathbf{U}_p}^\perp \boldsymbol{\mu}_b \right)^T. \end{aligned} \quad (22)$$

We obtain the OSP-based GLRT detector (OSP-GLRT), $\delta^{\text{OSP-GLRT}}(\mathbf{r})$ given by

$$\begin{aligned} \delta^{\text{OSP-GLRT}}(\mathbf{r}) &= \delta^{\text{GLRT}}(\tilde{\mathbf{r}}) = [\hat{\alpha}_{\text{MLE}}(\tilde{\mathbf{r}}) \tilde{\mathbf{t}}]^T \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{r}} \\ &= \frac{1}{\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p} \left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} P_{\mathbf{U}_p}^\perp (\mathbf{r} - \boldsymbol{\mu}_b) \right)^2 \end{aligned} \quad (23)$$

with $\hat{\alpha}_{\text{MLE}}(\tilde{\mathbf{r}}) = (\tilde{\mathbf{t}}^T \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{r}}) / (\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p)$, specifically, its SNR

$$\begin{aligned} \text{SNR}^{\text{OSP-GLRT}}(\mathbf{r}) &= \text{var}[\delta^{\text{OSP-GLRT}}(\mathbf{r}) | H_0] \\ &= (\hat{\alpha}_{\text{MLE}}(\tilde{\mathbf{r}}) \tilde{\mathbf{t}})^T \tilde{\mathbf{K}}^{-1} (\hat{\alpha}_{\text{MLE}}(\tilde{\mathbf{r}}) \tilde{\mathbf{t}}) \\ &= \frac{1}{\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p} \left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} P_{\mathbf{U}_p}^\perp (\mathbf{r} - \boldsymbol{\mu}_b) \right)^2 \end{aligned} \quad (24)$$

which is identical to $\delta^{\text{OSP-GLRT}}(\mathbf{r})$ in (23). Table I lists LRT-derived directors derived in this section.

It is worth noting that there are three major contributions of this section.

- 1) Even though there are some tutorial articles available in [20] and [23], they were derived from a signal processing perspective. For those working in the remote sensing community, these works may be either too complicated to comprehend or lack sufficient BKG knowledge to understand. This section formulates an HTD problem according to three familiar types of target pixel detection, pure pixel, subpixel, and mixed pixel and rederive LRT-based detectors developed in [20] and [23] by a simple standard detection theory.
- 2) The problem of subpixel target detection is formulated by a binary composite hypothesis testing problem with the target abundance fraction considered as a parameter to be detected.

- 3) The problem formulation of mixed-pixel target detection is new and can be solved by the well-known OSP technique.

V. SNR-BASED HTD

Since the probability density functions, $p_0(\mathbf{r})$ and $p_1(\mathbf{r})$ used in LRT in (6), GLRT in (16), and OSP-GLRT in (23) are generally unknown in many practical applications, the NP detector in (9) is not applicable.

Interestingly, it has been shown in Section IV that the three detectors, namely, LRT, GLRT, and OSP-GLRT, are identical to their corresponding SNR-based detectors, (7) (17), (24). This leads to a new criterion that can be used to design a detector $\delta(\mathbf{r})$. It is known as deflection, $D(\delta(\mathbf{r}))$ in [5] defined by

$$D(\delta(\mathbf{r})) = \frac{(E[\delta(\mathbf{r})|H_1] - E[\delta(\mathbf{r})|H_0])^2}{\text{var}[\delta(\mathbf{r})|H_0]} \quad (25)$$

where $E[\delta(\mathbf{r})|H_j]$ is expectation calculated by probability density function $p_j(\mathbf{r})$ under hypothesis H_j . Equation (25) indeed derives SNR-based detectors given by (7), (17), (24).

The rationale for using SNR can be found in the bottom paragraph [4, Sec. II-B, eq. (31)]. SNR turns out to be the parameter that determines false alarm P_F and detection probability, P_D with more details referred to [4, (Sec. III-B, eq.33), p. 56].

According to (25), this section develops a theory using SNR as an alternative to LRT, called SNR detection theory, to design detectors without appealing for probability density functions, $p_0(\mathbf{r})$ and $p_1(\mathbf{r})$ used in LRT, GLRT, and OSP-GLRT. More specifically, the SNR detection theory offers a practical approach that uses SNR via (25) as a criterion to rederive LRT, GLRT, and OSP-GLRT without using probability distributions assumed in (1), (11), and (18). Its idea uses SNR as a criterion for optimality to design a finite impulse response (FIR) filter specified by an L -dimensional weighting vector \mathbf{w} as follows:

$$\delta_{\mathbf{w}}(\mathbf{r}) = \mathbf{w}^T \mathbf{r}. \quad (26)$$

where the detector $\delta_{\mathbf{w}}(\mathbf{r})$ operating on the original data space \mathbf{X} with $\mathbf{r} \in \mathbf{X}$. Coincidentally, this approach is similar to that used to derive CEM in [10] and [11] which utilizes minimum variance (i.e., least-squares error) instead of SNR as an optimal criterion.

A. Full Pure Pixel Target Detection

According to [4], the role of \mathbf{K}^{-1} used in the LRT detectors specified by (6), (16), and (23) is designed to decorrelate BKG which is described by a Gaussian distribution with the covariance matrix \mathbf{K} . To mimic its functionality without using probability distributions, three data whitening methods are proposed for SNR-detectors to accomplish what \mathbf{K}^{-1} does for LRT/GLRT-derived detectors.

1) *DS Space*: The DS is a well-known technique, which has been widely used by independent component analysis (ICA) for blind source separation [2, Ch. 6], [30], [31] to remove from the original data space \mathbf{X} its first and second-order data statistics, i.e., the global mean $\boldsymbol{\mu}$ and global sample covariance matrix, \mathbf{K} through the whitening matrix specified by $\mathbf{K}^{-1/2}$.

The resulting DS-space is denoted by $\hat{\mathbf{X}}$, and data sample and target signal can be transformed by

$$\hat{\mathbf{r}} = \mathbf{K}^{-1/2}(\mathbf{r} - \boldsymbol{\mu}_b) \quad \text{and} \quad \hat{\mathbf{t}} = \mathbf{K}^{-1/2}(\mathbf{t} - \boldsymbol{\mu}_b) \quad (27)$$

respectively. After DS, the spherized data $\hat{\mathbf{X}}$ has zero mean and the identity matrix being its covariance matrix. Then the linear filter (26) is applied to $\hat{\mathbf{X}}$ to yield

$$\delta_{\mathbf{w}}(\hat{\mathbf{r}}) = \mathbf{w}^T \hat{\mathbf{r}} \quad (28)$$

and its SNR is calculated by (25) and given by

$$\text{SNR}(\delta_{\mathbf{w}}(\hat{\mathbf{r}})) = \frac{\mathbf{w}^T \hat{\mathbf{t}} (\mathbf{w}^T \hat{\mathbf{t}})^T}{[E(\mathbf{w}^T \hat{\mathbf{n}})^2]} = \frac{(\mathbf{w}^T \hat{\mathbf{t}})^2}{[\mathbf{w}^T \mathbf{w}]} = \sigma^{-2} (\bar{\mathbf{w}}^T \hat{\mathbf{t}}) \leq \hat{\mathbf{t}}^T \hat{\mathbf{t}} \quad (29)$$

where $\bar{\mathbf{w}} = \frac{\mathbf{w}}{\|\mathbf{w}\|}$ with $\|\mathbf{w}\|^2 = \mathbf{w}^T \mathbf{w}$, $\hat{\mathbf{n}}$ is the spherized noise with $\boldsymbol{\mu}_b = 0$ and variance $\sigma^2 = 1$. Specifically, the equality in (29) holds if and only if

$$\mathbf{w}^* = \kappa \hat{\mathbf{t}} \quad (30)$$

for any constant κ by Schwarz's inequality.

Substituting \mathbf{w}^* in (30) into (29) results in the maximal SNR given by

$$\text{SNR}(\delta_{\mathbf{w}^*}(\hat{\mathbf{r}})) = \max_{\mathbf{w}} \text{SNR}(\delta_{\mathbf{w}}(\hat{\mathbf{r}})) = \hat{\mathbf{t}}^T \hat{\mathbf{t}} \quad (31)$$

where the scalar constant κ vanishes since it is canceled out in (29).

Substituting (30) and (27) back to (28) and (31), we obtain an SNR-based detector operating on the spherized data $\hat{\mathbf{X}}$ by (27), which is defined by

$$\begin{aligned} \delta_{\kappa}^{\text{DS-SNR}}(\mathbf{r}) &= \kappa (\mathbf{K}^{-1/2}(\mathbf{t} - \boldsymbol{\mu}_b))^T \mathbf{K}^{-1/2}(\mathbf{r} - \boldsymbol{\mu}_b) \\ &= \kappa (\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1}(\mathbf{r} - \boldsymbol{\mu}_b) \end{aligned} \quad (32)$$

where DS used as a superscript in (32) indicates DS carried out by (27) before calculating SNR in (29). The maximum SNR of (32) can be found by

$$\begin{aligned} \text{SNR}(\delta_{\kappa}^{\text{DS-SNR}}(\mathbf{r})) &= \max_{\mathbf{w}} \text{SNR}(\delta_{\mathbf{w}}^{\text{DS-SNR}}(\mathbf{r})) \\ &= (\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1}(\mathbf{t} - \boldsymbol{\mu}_b) = \|\mathbf{K}^{-1/2}(\mathbf{t} - \boldsymbol{\mu}_b)\|^2. \end{aligned} \quad (33)$$

Comparing (32) to (6), there are two key differences between them. One is a constant κ in (32) which is absent in (6). A second difference is that \mathbf{t} in (32) removes the first-order statistics, $\boldsymbol{\mu}_b$, while \mathbf{t} in (6) does not. Since DS does not make any assumption on probability distributions, the covariance matrix \mathbf{K} is not necessarily a Gaussian covariance matrix as it is assumed in LRT/GLRT. Now, there are two scenarios to specify the constant κ for (32).

- 1) By setting κ in (32) to 1, then (32) becomes

$$\delta_{\kappa=1}^{\text{DS-SNR}}(\mathbf{r}) = (\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1}(\mathbf{r} - \boldsymbol{\mu}_b) = \delta^{\text{AMD}}(\mathbf{r}) \quad (34)$$

which a new detector, referred to as adaptive matched detector (AMD) in correspondence to the AMF in [14]. It is reduced to MD if the two signals, \mathbf{t} and \mathbf{r} are identical. In particular, when \mathbf{t} is replaced by \mathbf{r} in (34), AMD becomes RXD in [24].

- 2) Assume that the data sample \mathbf{r} is exactly \mathbf{t} . In this case, $\delta_{\kappa}^{\text{DS-SNR}}(\mathbf{t})$ will detect \mathbf{r} as \mathbf{t} with full abundance fraction, one. That is

$$\delta_{\kappa}^{\text{DS-SNR}}(\mathbf{t}) = 1 \Rightarrow \kappa = [(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)]^{-1}. \quad (35)$$

Substituting (35) into (34) yields the following normalized DS-SNR:

$$\delta^{\text{NDS-SNR}}(\mathbf{r}) = \frac{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)}{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b)} = \delta^{\text{NMAD}}(\mathbf{r}) \quad (36)$$

which becomes the normalized AMD (NAMD), which can be also referred to as normalized Mahalanobis distance (NMD). The process using (35) to derive (36) is generally called normalization.

2) *Covariance Matrix-Based Whitening*: The purpose of using (27) is designed to remove the first two orders of data statistics by centering the original data at the BKG mean, $\boldsymbol{\mu}_b = 0$ with all covariances made zero and variances made unity. Interestingly, if we are only interested in removing the second order of data statistics, while retaining the mean $\boldsymbol{\mu}_b$ by letting

$$\hat{\mathbf{r}} = \mathbf{K}^{-1/2} \mathbf{r}, \quad \hat{\mathbf{t}} = \mathbf{K}^{-1/2} \mathbf{t} \text{ and } \hat{\mathbf{b}} = \mathbf{K}^{-1/2} \mathbf{b} \quad (37)$$

such that $E[\hat{\mathbf{b}}\hat{\mathbf{b}}^T] = \sigma^2 \mathbf{I}$. Then the K-space is defined as the covariance matrix-based whitened data space, denoted by $\hat{\mathbf{X}}$, which is obtained by using the whitening matrix $\mathbf{K}^{-1/2}$ specified by (37). In this case, (26) becomes

$$\delta_{\mathbf{w}}(\hat{\mathbf{r}}) = \mathbf{w}^T \hat{\mathbf{r}} \quad (38)$$

and (25) can be calculated as

$$\delta_{\kappa}^{\text{K-SNR}}(\hat{\mathbf{r}}) = \kappa \hat{\mathbf{t}}^T \hat{\mathbf{r}} = \kappa (\mathbf{K}^{-1/2} \mathbf{t})^T (\mathbf{K}^{-1/2} \mathbf{t}) = \kappa \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r} \quad (39)$$

and

$$\max_{\hat{\mathbf{r}}} \text{SNR} \left[\delta_{\kappa}^{\text{K-SNR}}(\hat{\mathbf{r}}) \right] = \mathbf{t}^T \mathbf{K}^{-1} \mathbf{t} \quad (40)$$

where K-SNR is used in (39) to emphasize the fact that SNR is calculated in the K-space. Comparing (39) to (6), both $\delta_{\kappa}^{\text{K-SNR}}(\hat{\mathbf{r}})$ and $\delta^{\text{LRT}}(\mathbf{r})$ result in the same max SNR given by (40) and (7). With the same strategy used in (34) we can set $\kappa = 1$ to yield (6). On the other hand, we can also let

$$\delta_{\kappa}^{\text{K-SNR}}(\mathbf{t}) = 1 \Rightarrow \kappa = (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1}. \quad (41)$$

Substituting (41) back to (39) gives rise to

$$\delta^{\text{NK-SNR}}(\mathbf{r}) = \frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}} \equiv \delta^{\text{NLRT}}(\mathbf{r}) \quad (42)$$

which is called normalized K-SNR in (39) and can be also called normalized LRT (NLRT) of LRT in (6).

3) *Correlation Matrix-Based Whitening*: Furthermore, we also follow the same treatment given in Section V-A.2 by replacing \mathbf{K}^{-1} in (37) and (30) with the global sample correlation matrix \mathbf{R} to whiten the data space \mathbf{X} . Let the resulting whitened data space be defined as R-space, denoted by $\hat{\mathbf{X}}$ and the whitened data sample be denoted by

$$\hat{\mathbf{r}} = \mathbf{R}^{-1/2} \mathbf{r}. \quad (43)$$

The difference between DS-space $\hat{\mathbf{X}}$ in (27) and R-space $\hat{\mathbf{X}}$ in (43) is that (27) removes the first two orders of data statistics, mean and variance/covariance, while (43) only removes the second order of data statistics to make covariances zero and the noise variance has been normalized to σ^{-2} . As a result, $\hat{\mathbf{X}}$ does not center its global mean at the origin as $\hat{\mathbf{X}}$ does.

The linear filter (26) is then operated on $\hat{\mathbf{X}}$ to yield

$$\delta_{\mathbf{w}}(\hat{\mathbf{r}}) = \mathbf{w}^T \hat{\mathbf{r}} \quad (44)$$

and its SNR calculated by (25) in the R-space is given by

$$\text{SNR}(\delta_{\mathbf{w}}(\hat{\mathbf{r}})) = \frac{\mathbf{w}^T \hat{\mathbf{t}} (\mathbf{w}^T \hat{\mathbf{t}})^T}{\left[E(\mathbf{w}^T \hat{\mathbf{n}})^2 \right]} = \frac{(\mathbf{w}^T \hat{\mathbf{t}})^2}{[\mathbf{w}^T \mathbf{w}]} = \sigma^{-2} (\hat{\mathbf{w}}^T \hat{\mathbf{t}}) \leq \hat{\mathbf{t}}^T \hat{\mathbf{t}} \quad (45)$$

where $\hat{\mathbf{w}} = \frac{\mathbf{w}}{\|\mathbf{w}\|}$ with $\|\mathbf{w}\|^2 = \mathbf{w}^T \mathbf{w}$, $\hat{\mathbf{n}}$ is the whitened noise with variance σ^2 . Specifically, the equality in (45) holds if and only if

$$\mathbf{w}^* = \kappa \hat{\mathbf{t}} \quad (46)$$

for any constant κ by Schwarz's inequality.

Substituting \mathbf{w}^* in (46) into (45) results in the maximal SNR given by

$$\text{SNR}(\delta_{\mathbf{w}^*}(\hat{\mathbf{r}})) = \max_{\mathbf{w}} \text{SNR}(\delta_{\mathbf{w}}(\hat{\mathbf{r}})) = \hat{\mathbf{t}}^T \hat{\mathbf{t}} \quad (47)$$

where the scalar constant κ vanishes since it is canceled out in (45).

Substituting (46) back to (44) yields

$$\delta_{\kappa}^{\text{R-SNR}}(\hat{\mathbf{r}}) = \kappa \hat{\mathbf{t}}^T \hat{\mathbf{r}} = \kappa (\mathbf{R}^{-1/2} \mathbf{t})^T (\mathbf{R}^{-1/2} \mathbf{t}) = \kappa \mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}. \quad (48)$$

Once again, two scenarios can be derived from the parameter κ in (47).

1) By setting κ in (48) to 1, then (48) becomes

$$\delta_{\kappa=1}^{\text{R-SNR}}(\mathbf{r}) = \mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}. \quad (49)$$

where R-SNR is used in (49) to emphasize the fact that SNR in (48) is calculated in the R-space using (43) to obtain (49) along with its maximum SNR given by

$$\text{SNR}(\delta_{\kappa}^{\text{R-SNR}}(\hat{\mathbf{r}})) = \max_{\mathbf{w}} \text{SNR}(\delta_{\mathbf{w}}^{\text{R-SNR}}(\hat{\mathbf{r}})) = \mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}. \quad (50)$$

2) Now assume that (48) detects \mathbf{t} as its full abundance fraction as

$$\delta_{\kappa}^{\text{R-SNR}}(\mathbf{t}) = 1 \Rightarrow \kappa = (\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t})^{-1}. \quad (51)$$

Substituting (50) back to (48) results in a normalized R-SNR (NR-SNR)

$$\delta^{\text{NR-SNR}}(\mathbf{r}) = \frac{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}} \quad (52)$$

which turns out to be exactly the CEM derived in [10] and [11] which is described in the following.

4) *CEM*: As noted in [10], CEM also uses an FIR filter specified by (26) to solve the following linearly constrained optimization problem:

$$\min_{\mathbf{w}} \{ \mathbf{w}^T \mathbf{R} \mathbf{w} \} \text{ subject to } \mathbf{t}^T \mathbf{w} = \mathbf{w}^T \mathbf{t} = 1 \quad (53)$$

where $\mathbf{R} = (1/N) \sum_{i=1}^N \mathbf{r}_i \mathbf{r}_i^T$ is the auto-correlation sample matrix of $\{\mathbf{r}_i\}_{i=1}^N$. The optimal solution to (53), \mathbf{w}^{CEM} can be shown to be

$$\mathbf{w}^{\text{CEM}} = \frac{\mathbf{R}_{L \times L}^{-1} \mathbf{d}}{\mathbf{d}^T \mathbf{R}_{L \times L}^{-1} \mathbf{d}} \quad (54)$$

and CEM, denoted by $\delta^{\text{CEM}}(\mathbf{r})$ is given by

$$\delta^{\text{CEM}}(\mathbf{r}) = (\mathbf{w}^{\text{CEM}})^T \mathbf{r} = \frac{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}}. \quad (55)$$

Comparing (55) to (52), $\delta^{\text{NR-SNR}}(\mathbf{r})$ in (52) is exactly $\delta^{\text{CEM}}(\mathbf{r})$ in (55), which can be considered as an $\mathbf{R}^{-1/2}$ -whitened counterpart of NLRT in (42).

One note on CEM is particularly worthwhile. The constraint $\mathbf{w}^T \mathbf{t} = 1$ imposed on CEM in (53) is designed to lock in a signal arrival direction specified by the desired target signature \mathbf{t} through an array of sensors. This is equivalent to a target detector which detects a desired target pixel \mathbf{t} with full detected abundance fraction of 100% constrained by: 1) (35) to derive normalized data spheroid (NDS)-SNR in (36); 2) (41) to derive normalized K-whitened (NK)-SNR in (42); and 3) (51) to derive NR-SNR in (52). The criterion for optimality used by CEM is energy measured by least-squares error which also turns out to be minimum variance in (25) used by SNR. Therefore, it is not surprising to see that CEM is indeed the NR-SNR. Interestingly, the process of imposing this constraint in HTD is generally called normalization without giving any reason such as NAMD (i.e., NMD) in (36), NAMF in [16], and [17]. A similar observation is also made on OSP [32] where *a posteriori* OSP developed in [33] and [34] is actually obtained by imposing the constraint (51) on the OSP in [6] because OSP is also an SNR-derived target detector. This explains that OSP originally designed as a target detector can be also converted to *a posteriori* OSP to be used as a linear spectral unmixing technique to unmix pixels for their abundance fractions [35].

B. Subpixel Target Detection

Since no probability distributions are available to find MLE, we use MD instead of MLE as a measure to calculate the estimation error defined by

$$\hat{\alpha}_{\text{MD}}(\mathbf{r}) = \arg \left\{ \min_{\alpha} (\mathbf{r} - \alpha \mathbf{t})^T \mathbf{K}^{-1} (\mathbf{r} - \alpha \mathbf{t}) \right\} \quad (56)$$

which turns out to be the same as MLE in (15).

1) *DS-Space*: Now, if we use (27) again, then (32) can be extended to GSNR by including (56) in SNR which is defined and given by

$$\text{GDS-SNR}(\delta_{\mathbf{w}}(\hat{\mathbf{r}})) = \frac{\mathbf{w}^T \hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}} (\mathbf{w}^T \hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}})^T}{\left[E(\mathbf{w}^T \hat{\mathbf{b}})^2 \right]} \quad (62)$$

$$= \frac{(\mathbf{w}^T \hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}})^2}{[\mathbf{w}^T \mathbf{w}]} \leq (\hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}})^T \hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}} \quad (57)$$

where $\bar{\mathbf{w}} = (\mathbf{w} / \|\mathbf{w}\|)$ and the equality holds if and only if

$$\mathbf{w}^*(\hat{\mathbf{r}}) = \kappa \hat{\alpha}_{\text{MD}}(\hat{\mathbf{r}}) \hat{\mathbf{t}}. \quad (58)$$

Therefore, by substituting (58) into (57), we obtain the following optimal detector:

$$\begin{aligned} \delta_{\kappa}^{\text{GDS-SNR}}(\hat{\mathbf{r}}) &= (\mathbf{w}^*)^T \hat{\mathbf{r}} \\ &= \kappa (\hat{\alpha}_{\text{MD}}(\mathbf{r}) \mathbf{K}^{-1/2} (\mathbf{t} - \mu_b))^T (\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{r} - \mu_b) \\ &= \kappa \frac{(\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{r} - \mu_b)}{(\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{t} - \mu_b)} (\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{r} - \mu_b) \\ &= \kappa \frac{((\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{r} - \mu_b))^2}{(\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{t} - \mu_b)} \end{aligned} \quad (59)$$

where the subscript of “GDS-SNR” is used to emphasize the SNR is calculated by using (56) in the DS-space, $\hat{\mathbf{X}}$.

Similar to DS-SNR, two scenarios can be derived from the parameter κ in (59)

1) By setting κ in (59) to 1, then (59) becomes

$$\delta_{\kappa=1}^{\text{GDS-SNR}}(\mathbf{r}) = \frac{((\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{r} - \mu_b))^2}{(\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{t} - \mu_b)}. \quad (60)$$

2) By setting $\delta_{\kappa}^{\text{GDS-SNR}}(\mathbf{t}) = 1$ in (59), we can derive $\kappa = ((\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{t} - \mu_b))^{-1}$ and

$$\begin{aligned} \delta^{\text{NGDS-SNR}}(\mathbf{r}) &= \left(\frac{((\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{r} - \mu_b))^2}{(\mathbf{t} - \mu_b)^T \mathbf{K}^{-1} (\mathbf{t} - \mu_b)} \right)^2 \\ &= (\delta^{\text{NMAD}}(\mathbf{r}))^2 \end{aligned} \quad (61)$$

which is called NGK-SNR. It also turns out to be the square of NAMD in (36).

2) *K-Space*: Alternatively, if we use (37) and (38), then we can extend the SNR of $\delta_{\mathbf{w}}(\hat{\mathbf{r}})$ in (40) to GSNR by including (56) in SNR which is defined and given by

$$\begin{aligned} \text{GK-SNR}(\delta_{\mathbf{w}}(\hat{\mathbf{r}})) &= \frac{\mathbf{w}^T \hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}} (\mathbf{w}^T \hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}})^T}{\left[E(\mathbf{w}^T \hat{\mathbf{b}})^2 \right]} \\ &= \frac{(\mathbf{w}^T \hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}})^2}{[\mathbf{w}^T \mathbf{w}]} = \sigma^{-2} \left| \bar{\mathbf{w}}^T \hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}} \right|^2 \\ &\leq \left(\hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}} \right)^T \hat{\alpha}_{\text{MD}}(\mathbf{r}) \hat{\mathbf{t}} \end{aligned} \quad (62)$$

where $\bar{\mathbf{w}} = (\mathbf{w} / \|\mathbf{w}\|)$ and the equality holds if and only if

$$\mathbf{w}^*(\hat{\mathbf{r}}) = \kappa \hat{\alpha}_{\text{MD}}(\hat{\mathbf{r}}) \hat{\mathbf{t}}. \quad (63)$$

Similarly, GK-SNR is used in (62) to emphasize the fact that GSNR is calculated on the $\mathbf{K}^{-1/2}$ -whitened space by (37)

TABLE II
LIST OF SNR-DERIVED DETECTORS

SNR-derived detection methods	Subpixel/mixed pixel detectors	Equation
DS-SNR in DS-space	$\delta_{\kappa}^{\text{DS-SNR}}(\mathbf{r}) = \kappa (\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)$	(32)
K-SNR in K-space	$\delta_{\kappa}^{\text{K-SNR}}(\mathbf{r}) = \kappa \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}$	(39)
R-SNR in R-space	$\delta_{\kappa}^{\text{R-SNR}}(\mathbf{r}) = \kappa \mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}$	(48)
GDS-SNR: GSNR in DS-space	$\delta_{\kappa}^{\text{GDS-SNR}}(\mathbf{r}) = \kappa \frac{\left((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b) \right)^2}{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b)}$	(59)
GK-SNR: GSNR in K-space	$\delta_{\kappa}^{\text{GK-SNR}}(\mathbf{r}) = \kappa \frac{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}}$	(64)
GR-SNR: GSNR in R-space	$\delta_{\kappa}^{\text{GR-SNR}}(\mathbf{r}) = \kappa \frac{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r})^2}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}}$	(67)
OSP-GK-SNR: GSNR in K-space	$\delta_{\kappa}^{\text{OSP-GK-SNR}}(\mathbf{r}) = \kappa \frac{\left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} P_{\mathbf{U}_p}^T \mathbf{r} \right)^2}{\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p}$	(72)

to obtain (62). So, by substituting (63) into (64), we obtain the following optimal detector

$$\begin{aligned} \delta_{\kappa}^{\text{GK-SNR}}(\mathbf{r}) &= (\mathbf{w}^*)^T \tilde{\mathbf{r}} = \kappa (\hat{\alpha}_{\text{MD}}(\mathbf{r}) \mathbf{K}^{-1/2} \mathbf{t})^T \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r} \\ &= \kappa \frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}} \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r} = \kappa \frac{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}}. \end{aligned} \quad (64)$$

There are also two scenarios resulting from setting particular values for the κ in (64).

1) Setting $\kappa = 1$ yields $\delta_{\kappa=1}^{\text{GK-SNR}}(\mathbf{r}) = ((\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2 / \mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})$ which turns out to be exactly AMF in (16)

2) Since $\delta_{\kappa}^{\text{GK-SNR}}(\mathbf{t}) = 1$, the κ in (60) can be obtained by

$$\kappa = (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1}. \quad (65)$$

Substituting (65) into (64) yields NGK-SNR

$$\delta_{\kappa=(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1}}^{\text{NGK-SNR}}(\mathbf{r}) = \left(\frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}} \right)^2 = \delta^{\text{ASD}}(\mathbf{r}) = \delta^{\text{NAMF}}(\mathbf{r}) \quad (66)$$

which is exactly the ASD derived in [16] and [17] and can be also considered as normalized version of AMF in (16), (NAMF).

3) *R-Space*: In analogy with (62), we can also derive GSNR by operating data sample $\tilde{\mathbf{r}} = \mathbf{R}^{-1/2} \mathbf{r}$ in the R-space using (43) as follows:

$$\delta_{\kappa}^{\text{GR-SNR}}(\mathbf{r}) = \kappa \frac{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r})^2}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}} \quad (67)$$

and

$$\delta_{\kappa=1}^{\text{GR-SNR}}(\mathbf{r}) = \frac{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r})^2}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}} \quad (68)$$

where GR-SNR is used in (67) to emphasize the fact that SNR is calculated on the $\mathbf{R}^{-1/2}$ -whitened space by (43) to

obtain (67). Now, using $\delta_{\kappa}^{\text{GR-SNR}}(\mathbf{t}) = 1$ derives

$$\delta_{\kappa=(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t})^{-1}}^{\text{GR-SNR}}(\mathbf{r}) = \left(\frac{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}} \right)^2 = \delta^{\text{NGR-SNR}}(\mathbf{r}) \quad (69)$$

which is defined as NGR-SNR and occurs to be the squared CEM.

C. Mixed Target Detection

Following the same treatment in Section IV-C an OSP-GK-SNR detector, $\delta_{\kappa}^{\text{OSP-GK-SNR}}(\mathbf{r})$ can be derived from (64) as

$$\delta_{\kappa}^{\text{OSP-GK-SNR}}(\tilde{\mathbf{r}}) = [\hat{\alpha}_{\text{MD}}(\tilde{\mathbf{r}}) \tilde{\mathbf{t}}]^T \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{r}} = \kappa \frac{(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{r}})^2}{\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p} \quad (70)$$

and

$$\delta_{\kappa}^{\text{OSP-GK-SNR}}(\tilde{\mathbf{t}}) = 1 \Rightarrow \kappa = (\tilde{\mathbf{t}}^T \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{t}})^{-1}. \quad (71)$$

Substituting (71) into (70) yields a NOSP-GK-SNR defined by

$$\begin{aligned} \delta^{\text{NOSP-GK-SNR}}(\mathbf{r}) &= [\hat{\alpha}_{\text{MD}}(\tilde{\mathbf{r}}) \tilde{\mathbf{t}}]^T \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{r}} = \left(\frac{\tilde{\mathbf{t}}^T \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{r}}}{\tilde{\mathbf{t}}^T \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{t}}} \right)^2 \\ &= \left[\frac{\left(P_{\mathbf{U}_p}^{\perp} \mathbf{t} \right)^T \tilde{\mathbf{K}}^{-1} \left(P_{\mathbf{U}_p}^{\perp} (\mathbf{r} - \boldsymbol{\mu}_b) \right)}{\left(P_{\mathbf{U}_p}^{\perp} \mathbf{t} \right)^T \tilde{\mathbf{K}}^{-1} \left(P_{\mathbf{U}_p}^{\perp} \mathbf{t} \right)} \right]^2. \end{aligned} \quad (72)$$

Finally, Table II summarizes three versions of detectors that uses (25) as an SNR criterion and are operated in three types of data spaces, DS-space by (27), K-space by (37), and R-space by (43).

The major contributions of this section are summarized in the following.

- 1) The entire SNR-based detection theory section is new and has never been explored in the past.

- 2) The SNR-based detection theory is derived from deflection criterion based on SNR without appealing for hypothesis testing formulation. As a result, no probability distributions are needed.
- 3) The SNR-based detection theory introduces three data whitening techniques to include many existing LRT-derived detectors as special cases.
- 4) The SNR-based detection theory integrates and unifies all LRT-derived detectors derived in the literature under its umbrella.
- 5) Several new SNR-based detectors are derived and have never been reported in the past.
- 6) The SNR-based detection theory also provides rationales on why many existing detectors are obtained by normalization.

VI. RELATIONSHIP BETWEEN LRT-DERIVED DETECTORS AND SNR-DERIVED DETECTORS

The relationship between LRT-derived detectors in Section IV and SNR-derived detectors in Section V is established by deflection criterion specified by (25) which is essentially SNR. There are several key differences between them. One is their problem formulations. LRT is the optimal filter resulting from a hypothesis testing problem specified by (2), (11) or (18). Compared to LRT, SNR-based detectors are derived from linear filters specified by (26) using SNR as an optimal criterion. Another is the availability of the probability distribution under each hypothesis is assumed so that LRT can be calculated, particularly, Gaussian distribution assumption to derive analytic forms for many various LRT detectors. By contrast, the SNR-based approach requires only the first two orders of data statistics without assuming probability distributions. Most importantly, the threshold τ used in LRT (3) as a parameter must be determined by either costs and prior probability distribution or a constraint β imposed on P_F formulated by hypothesis testing problems in correspondence to κ used in SNR which can be determined by applications. Indeed, it is the parameter κ which makes SNR-approach more interesting and appealing than the LRT approach. In order to see that, Table III provides a comparison of one-to-one correspondence between SNR-derived detectors and LRT-derived detectors where the parameter κ is set to two scenarios. One is $\kappa = 1$ which indicates the SNR-derived detectors are free of constant. The other is to set κ to the value that make the abundance fraction of the matched target pixel unity, i.e., $\delta_w(t) = w^T t = 1$, which is exactly the same constraint used in (53) to derive CEM. Such process of setting $\delta_w(t) = w^T t = 1$ is generally referred to as normalization which is commonly used in HTD but is never explained why it is called normalization. As a consequence, LRT-derived detectors can be reinterpreted under the umbrella of the SNR-detection theory.

As shown in Table III, most of LRT detectors derived in the literature can find their counterparts corresponding to SNR-derived detectors. Also, by manipulating the parameter κ the LRT-based detectors in Table I can be also considered as SNR/GSNR, $\delta_{\kappa}^{\text{SNR/GSNR}}(r)$ in Table III. Specifically, Table III can be used to expand Tables I-II to Table IV

by including their corresponding normalized counterparts obtained by imposing the constraint on the parameter κ , i.e., $\delta_{\kappa}^{\text{SNR/GSNR}}(t) = 1$ in the same manner that the constraint, $t^T w = w^T t = 1$ imposed on CEM.

It is interesting to note that we can take advantage of (19) to define the OSP projection matrix P_U on any matrix U as

$$P_U = UU^{\#} = U(U^T U)^{-1} U^T \quad (73)$$

so that $P_U^{\perp} = I - P_U$. By means of (73), we further rederive the following equivalent forms for various LRT- and SNR-based detectors derived in Sections III and IV as follows:

$$\begin{aligned} & \delta_{\kappa=1}^{\text{GK-SNR}}(r) \\ &= \delta^{\text{AMF}}(r) = \delta^{\text{GLRT}}(r) = \frac{(t^T K^{-1} r)^2}{t^T K^{-1} t} \\ &= r^T K^{-1} t (t^T K^{-1} t)^{-1} (r^T K^{-1} t)^T \\ &= (K^{-1/2} r)^T (K^{-1/2} t) ((K^{-1/2} t)^T (K^{-1/2} t))^{-1} (K^{-1/2} t)^T (K^{-1/2} r) \\ &= (K^{-1/2} r)^T P_{K^{-1/2} t} (K^{-1/2} r) \text{ by (73)} \\ &= \|P_{K^{-1/2} t}(K^{-1/2} r)\|^2 \text{ because of } P_{K^{-1/2} t}^2 = P_{K^{-1/2} t} \\ &= \|P_{\hat{t}}(\hat{r})\|^2 \text{ in the K-space via (37)} \end{aligned} \quad (74)$$

along with its normalized version given by

$$\begin{aligned} \delta^{\text{ASD}}(r) &= \delta^{\text{NAMF}}(r) = \delta^{\text{NGLRT}}(r) \\ &= \left(\frac{t^T K^{-1} r}{t^T K^{-1} t} \right)^2 = \frac{\|P_{K^{-1/2} t}(K^{-1/2} r)\|^2}{t^T K^{-1} t} \\ &= \frac{\|P_{\hat{t}}(\hat{r})\|^2}{\|\hat{t}\|^2} \text{ in the K-space.} \end{aligned} \quad (75)$$

Similarly, using (73) and (27) to sphere the data space we can derive the following detectors in the DS-space:

$$\begin{aligned} \delta_{\kappa=1}^{\text{GDS-SNR}}(r) &= \frac{((t - \mu_b)^T K^{-1} (r - \mu_b))^2}{(t - \mu_b)^T K^{-1} (t - \mu_b)} \\ &= \|P_{K^{-1/2}(t-\mu_b)}(K^{-1/2}(r - \mu_b))\|^2 \\ &= \|P_{\hat{t}}(\hat{r})\|^2 \text{ in the DS-space via (27)} \end{aligned} \quad (76)$$

along with its normalized version

$$\begin{aligned} \delta^{\text{NAMD}}(r) &= \frac{\|P_{K^{-1/2}(t-\mu_b)}(K^{-1/2}(r - \mu_b))\|^2}{(t - \mu_b)^T K^{-1} (t - \mu_b)} \\ &= \frac{\|P_{\hat{t}}(\hat{r})\|^2}{\|\hat{t}\|^2} \text{ in the DS-space via (27).} \end{aligned} \quad (77)$$

Furthermore, using (73) and (43) to whiten the data space by $R^{-1/2}$ we can derive the following detectors in the R-space:

$$\begin{aligned} \delta_{\kappa=1}^{\text{GR-SNR}}(r) &= \frac{(t^T R^{-1} r)^2}{t^T R^{-1} t} = \|P_{R^{-1/2} t}(R^{-1/2} r)\|^2 \\ &= \|P_{\hat{t}}(\hat{r})\|^2 \text{ in the R-space via (43)} \end{aligned} \quad (78)$$

along with its normalized version

$$(\delta^{\text{CEM}}(r))^2 = \frac{\|P_{R^{-1/2} t}(R^{-1/2} r)\|^2}{t^T R^{-1} t}$$

TABLE III
RELATIONSHIP BETWEEN LRT-DERIVED DETECTORS DERIVED FROM SNR-DERIVED DETECTORS

SNR-based detection methods	κ	LRT-Mixed pixel detectors	equation
DS-SNR detector $\delta_{\kappa}^{\text{DS-SNR}}(\mathbf{r}) = \kappa (\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)$	$\kappa=1$ $\delta^{\text{AMD}}(\mathbf{r}) = (\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)$ $\kappa = \left((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b) \right)^{-1}$ $\delta^{\text{NAMD}}(\mathbf{r}) = \frac{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)}{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b)}$	NA	(32) (36)
K-SNR detector $\delta_{\kappa}^{\text{K-SNR}}(\mathbf{r}) = \kappa \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}$	$\kappa=1$ $\delta^{\text{LRT}}(\mathbf{r}) = \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}$ $\kappa = (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1}$ $\delta^{\text{NK-SNR}}(\mathbf{r}) = \frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}} = \delta^{\text{NLRT}}(\mathbf{r})$	NA	(6) (42)
R-SNR detector $\delta_{\kappa}^{\text{R-SNR}}(\mathbf{r}) = \kappa \mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}$	$\kappa=1$ $\delta_{\kappa=1}^{\text{R-SNR}}(\mathbf{r}) = \mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}$ $\kappa = (\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t})^{-1}$ $\delta^{\text{NR-SNR}}(\mathbf{r}) = \frac{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}} = \delta^{\text{CEM}}(\mathbf{r})$	NA	(49) (52)
GDS-SNR detector $\kappa \frac{((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b))^2}{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b)}$	$\kappa=1$ $\frac{((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b))^2}{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b)}$ $\kappa = \left((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b) \right)^{-1}$ $\left(\frac{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)}{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b)} \right)^2 = \left(\delta^{\text{NAMD}}(\mathbf{r}) \right)^2$		(60) (61)
GK-SNR detector $\delta_{\kappa}^{\text{GK-SNR}}(\mathbf{r}) = \kappa \frac{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}}$	$\kappa=1$ $\delta^{\text{GLRT}}(\mathbf{r}) = \left(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t} \right)^{-1} \left(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r} \right)^2$ $\kappa = (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1}$ $\delta^{\text{NGK-SNR}}(\mathbf{r}) = \left(\frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}} \right)^2 = \delta^{\text{ASD}}(\mathbf{r})$	NA	(16) (66)
GR-SNR detector $\delta_{\kappa}^{\text{GR-SNR}}(\mathbf{r}) = \kappa \frac{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r})^2}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}}$	$\kappa=1$ $\delta_{\kappa=1}^{\text{GR-SNR}}(\mathbf{r}) = \left(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t} \right)^{-1} \left(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r} \right)^2$ $\kappa = (\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t})^{-1}$ $\delta^{\text{NGR-SNR}}(\mathbf{r}) = \left(\frac{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}} \right)^2 = \left(\delta^{\text{CEM}}(\mathbf{r}) \right)^2$	NA	(68) (69)
OSP-GSNR $\delta_{\kappa}^{\text{OSP-GSNR}}(\mathbf{r}) = \kappa \frac{\left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} P_{\mathbf{U}_p}^T \mathbf{r} \right)^2}{\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p}$	NA	$\tilde{\delta}^{\text{GLRT}}(\mathbf{r}) = \left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p \right)^{-1} \left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} P_{\mathbf{U}_p}^T (\mathbf{r} - \boldsymbol{\mu}_b) \right)^2$ $\kappa = (\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p)^{-1}$ $\delta^{\text{NOSP-GK-SNR}}(\mathbf{r}) = \left(\frac{\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} P_{\mathbf{U}_p}^T (\mathbf{r} - \boldsymbol{\mu}_b)}{\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p} \right)^2$	(23)=(70) (72)

TABLE IV
LIST OF TARGET DETECTORS DERIVED FROM LRT-BASED DETECTORS ALONG WITH THEIR NORMALIZED COUNTERPARTS

methods	Full pure pixel detectors	Subpixel detectors	Mixed pixel detectors	Equation
LRT	$\delta^{\text{LRT}}(\mathbf{r}) = \mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}$	NA	NA	(6)
NLRT $\kappa = (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1}$	$\delta^{\text{NLRT}}(\mathbf{r}) = \frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}}$	NA		(42)
GLRT=AMF	NA	$\delta^{\text{GLRT}}(\mathbf{r}) = (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1} (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2$ $= \delta^{\text{AMF}}(\mathbf{r})$	NA	(16)
NGLRT=NAMF $=\text{ASD}=(\text{NLRT})^2$ $\kappa = (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1}$	NA	$\delta^{\text{NGLRT}}(\mathbf{r}) = (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-2} (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2$ $= \delta^{\text{ASD}}(\mathbf{r}) = (\delta^{\text{NLRT}}(\mathbf{r}))^2$	NA	(66)
OSP-GLRT	NA	NA	$\delta^{\text{OSP-GLRT}}(\mathbf{r}) = (\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p)^{-1} (\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} R_{\mathbf{U}_p}^T (\mathbf{r} - \boldsymbol{\mu}_b))^2$	(23)
NOSP-GLRT $\kappa = (\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p)^{-1}$	NA	NA	$\delta^{\text{NOSP-GLRT}}(\mathbf{r}) = (\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p)^{-2} (\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} R_{\mathbf{U}_p}^T (\mathbf{r} - \boldsymbol{\mu}_b))^2$	(72)

$$= \frac{\|P_{\mathbf{t}}(\tilde{\mathbf{r}})\|^2}{\|\tilde{\mathbf{t}}\|^2} \text{ in the R-space via (43).} \quad (79)$$

By virtue of (73)–(79), many LRT-derived detectors, such GLRT, AMF, NAMD, ASD, etc., can be further shown to be identical to $\delta_{\kappa=1}^{\text{GK-SNR}}(\mathbf{r})$, $\delta_{\kappa=1}^{\text{GDS-SNR}}(\mathbf{r})$, $\delta_{\kappa=1}^{\text{GR-SNR}}(\mathbf{r})$ or their normalized versions, $\delta_{\kappa=(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1}}^{\text{GK-SNR}}(\mathbf{r})$, $\delta_{\kappa=((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b))^{-1}}^{\text{GDS-SNR}}(\mathbf{r})$, $\delta_{\kappa=(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t})^{-1}}^{\text{GR-SNR}}(\mathbf{r})$. In particular, these detectors are either OSP of data sample \mathbf{r} or its normalized OSP in three different processed spaces, DS-space by (27), K-space whitened by $\mathbf{K}^{-1/2}$ by (37) and the R-space whitened by $\mathbf{R}^{-1/2}$ by (43). These derivations demonstrate once again that LRT-derived detector can be considered as special cases of SNR-derived detectors.

The major contributions of this section are

- 1) Exploration of interrelationships between LRT-derived and SNR-derived detectors.
- 2) Reinterpretation of LRT-derived detector in terms of OSP in three data whitened spaces.

VII. SA-BASED DETECTORS

In Sections III and IV, the detectors are derived from LRT using Gaussian probability distributions and the SNR-based linear filter specified by (26), respectively. The resulting optimal filters from both approaches turn out to be linear MFs. However, on some occasions such a linear filter design may not be the best approach. Given the nature of an MF derived in Sections III and IV, the simplest nonlinear filter is to measure the angle between two spectral vectors, referred to as SA which performs a similar matching signal task, so-called SA mapper (SAM) [1]. However, due to sensitivity of SA to the BKG effect, directly applying SA to data samples in the original data space does not always work effectively. As noted in Sections III and IV, BKG can be generally characterized by the second-order data statistics. Accordingly, removing the

second-order statistics should be able to suppress BKG impact on SA so as to improve SA performance. In doing so, the three data whitening methods introduced in Sections III and IV, DS by $\mathbf{K}^{-1/2}(\mathbf{r} - \boldsymbol{\mu}_b)$ in (27) and data whitening by $\mathbf{K}^{-1/2}$ in (37) and $\mathbf{R}^{-1/2}$ in (43) can be used for this purpose.

First of all, we use (37) for data whitening to produce the K-space on which SA is operated. The SA is calculated between $\hat{\mathbf{t}}$ and $\hat{\mathbf{r}}$, denoted by $\hat{\theta}$ by

$$\cos \hat{\theta} = \frac{\langle \hat{\mathbf{t}}, \hat{\mathbf{r}} \rangle}{\|\hat{\mathbf{t}}\| \|\hat{\mathbf{r}}\|}. \quad (80)$$

Substituting (37) back to (80) gives rise to the normalized matched filter (NMF) defined in [15] and [21] as follows:

$$\delta^{\text{NMF}}(\mathbf{r}) = \frac{\langle \hat{\mathbf{t}}, \hat{\mathbf{r}} \rangle}{\|\hat{\mathbf{t}}\| \|\hat{\mathbf{r}}\|} = \frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{1/2} (\mathbf{r}^T \mathbf{K}^{-1} \mathbf{r})^{1/2}} \\ = \cos \hat{\theta} \equiv \delta^{\text{K-SA}}(\mathbf{r}) \quad (81)$$

which is more appropriately defined as K-SA than NMF since it uses $\mathbf{K}^{-1/2}$ in (37) to whiten the data space \mathbf{X} as to produce K-space, $\hat{\mathbf{X}}$ on which SA can be calculated by (80). Furthermore, taking the square of $\delta^{\text{K-SA}}(\mathbf{r})$ in (81) yields

$$\delta^{\text{K-SA}^2}(\mathbf{r}) \equiv (\delta^{\text{K-SA}}(\mathbf{r}))^2 = \frac{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2}{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}) (\mathbf{r}^T \mathbf{K}^{-1} \mathbf{r})} = \cos^2 \hat{\theta} \\ \Leftrightarrow \delta^{\text{ACE}}(\mathbf{r}) \equiv \delta^{\text{K-SA}^2}(\mathbf{r}) \quad (82)$$

which can be defined as K-SA². This is because it is operated on $\hat{\mathbf{X}}$ to square (82). It turns out that $\delta^{\text{K-SA}^2}(\mathbf{r})$ in (82) is exactly the well-known adaptive cosine estimate (ACE) [12], [16], [18], denoted by $\delta^{\text{ACE}}(\mathbf{r})$.

Furthermore, we can also use (27) instead of (37) to calculate the SA between $\hat{\mathbf{t}}$ and $\hat{\mathbf{r}}$ in $\hat{\mathbf{X}}$, denoted by $\hat{\theta}$

$$\begin{aligned} \cos \hat{\theta} &= \frac{\langle \hat{\mathbf{t}}, \hat{\mathbf{r}} \rangle}{\|\hat{\mathbf{t}}\| \|\hat{\mathbf{r}}\|} \\ &= \frac{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)}{((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b))^{1/2} ((\mathbf{r} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b))^{1/2}} \\ \Rightarrow \delta^{\text{DS-SA}^2}(\mathbf{r}) &= \cos^2 \hat{\theta} \end{aligned} \quad (83)$$

which is referred to as DS-SA² because SA is operated on the DS-space $\hat{\mathbf{X}}$.

Now, we can follow the same treatment used to derive (83) by replacing $\mathbf{K}^{-1/2}$ in (37) with $\mathbf{R}^{-1/2}$ in (43). Analogous to (83), we can calculate the angle between $\bar{\mathbf{t}}$ and $\bar{\mathbf{r}}$ in $\bar{\mathbf{X}}$, denoted by $\bar{\theta}$

$$\cos \bar{\theta} = \frac{\langle \bar{\mathbf{t}}, \bar{\mathbf{r}} \rangle}{\|\bar{\mathbf{t}}\| \|\bar{\mathbf{r}}\|} = \frac{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}}{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t})^{1/2} (\mathbf{r}^T \mathbf{R}^{-1} \mathbf{r})^{1/2}}. \quad (84)$$

Using (84) we further define a new detector similar to DS-SA² as

$$\delta^{\text{R-SA}^2}(\mathbf{r}) = \cos^2 \bar{\theta} = \frac{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r})^2}{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}) (\mathbf{r}^T \mathbf{R}^{-1} \mathbf{r})} \quad (85)$$

referred to as R-SA², where R-SA² is operated on the R-space, $\bar{\mathbf{X}}$ by squaring (84). Table IV lists the four data-whitened SA-based detectors derived in this section.

It should be noted the two detectors, NMF = K-SA and ACE = K-SA² in Table V are SA detectors because NMF and ACE are obtained by normalizing MF in [23] and MSD in [15], respectively, both of which involves SA between \mathbf{t} and \mathbf{r} .

The major contributions of this section can be described as follows.

- 1) The SA-based theory derived on three data whitened spaces, i.e., DS space, K-space, and R-space is new.
- 2) Two new SA-detectors, DS-SA², and R-SA² are new.
- 3) In particular, two new SA detectors, K-SA and K-SA² which can be interpreted as the commonly used AMF and ACE, respectively.

VIII. 3-D ROC CURVE-DERIVED DETECTION MEASURES

To further conduct a more detailed quantitative analysis on detection performance, a recent work in [9] proposed an effective 3-D ROC analysis-based evaluation tool, which extends the traditional 2-D ROC analysis by including the threshold τ used in the NP detector specified by (9) as an independent parameter. The idea of developing a 2-D ROC curve of P_D versus P_F is because the desired threshold τ is determined by a specific P_F imposed on a detector. The ROC curve is plotted by varying values of P_F from 0 to 1. Therefore, to see the effectiveness of a detector the area under an ROC curve (AUC), denoted by $\text{AUC}_{(D,F)}$, is calculated by the area under the 2-D ROC curve as an index to measure the effectiveness of a detector. However, there are two major issues in using $\text{AUC}_{(D,F)}$. One is that both P_D and P_F are calculated by the same threshold τ . Accordingly, when both P_D and P_F are very high, its calculated $\text{AUC}_{(D,F)}$ is also

very high. As we have seen in the literature, many results reported their $\text{AUC}_{(D,F)}$ values higher than those calculated by other detectors and concluded that their detectors were better than these detectors. This conclusion is misleading because $\text{AUC}_{(D,F)}$ does not reflect BKG suppressibility as demonstrated in [9] and the experiments conducted in Section IX. The development of 3-D ROC analysis arises from this need. It considers P_D , P_F , and τ as three independent parameters. As a result, a 3-D ROC curve can be generated by a triplet parameter vector specified by (P_D, P_F, τ) . From this 3-D ROC curve, three 2-D ROC curves of (P_D, P_F) , (P_D, τ) , (P_F, τ) can be also generated along with their respective AUC values, denoted by, $\text{AUC}_{(D,F)}$, $\text{AUC}_{(D,\tau)}$, and $\text{AUC}_{(F,\tau)}$, where the 2-D ROC curves of (P_D, P_F) becomes the commonly used 2-D ROC curve, $\text{AUC}_{(D,\tau)}$ and $\text{AUC}_{(F,\tau)}$ can be used to evaluate detection power and BKG suppressibility, respectively. By virtue of $\text{AUC}_{(D,F)}$, $\text{AUC}_{(D,\tau)}$, and $\text{AUC}_{(F,\tau)}$ generated from the 2-D ROC curves of (P_D, P_F) , (P_D, τ) , (P_F, τ) , we can further define the following detection measures. Target detectability (TD), denoted by AUC_{TD} and defined by

$$0 \leq \text{AUC}_{\text{TD}} = \text{AUC}_{(D,F)} + \text{AUC}_{(D,\tau)} \leq 2. \quad (86)$$

BS measure, denoted by AUC_{BS} and defined by

$$-1 \leq \text{AUC}_{\text{BS}} = \text{AUC}_{(D,F)} - \text{AUC}_{(F,\tau)} \leq 1. \quad (87)$$

Target detection and BS (TDBS), denoted by AUC_{TDBS} and defined by

$$-1 \leq \text{AUC}_{\text{TDBS}} = \text{AUC}_{(D,\tau)} - \text{AUC}_{(F,\tau)} \leq 1. \quad (88)$$

ODP, denoted by AUC_{ODP} and defined by

$$-1 \leq \text{AUC}_{\text{ODP}} = \text{AUC}_{(D,F)} + \text{AUC}_{(D,\tau)} - \text{AUC}_{(F,\tau)} \leq 2. \quad (89)$$

SNPR, denoted by AUC_{SNPR} and defined by

$$0 \leq \text{AUC}_{\text{SNPR}} = \frac{\text{AUC}_{(D,\tau)}}{\text{AUC}_{(F,\tau)}}. \quad (90)$$

The aforementioned detection measures were then used to evaluate the quantitative performance of detectors for experiments conducted in Section IX. More details can be found in [9] along with its software available on the website <https://wiki.umbc.edu/display/rssipl/10.+Download>.

IX. EXPERIMENTS

Table III in Section VI summarizes various SNR-derived detectors with their corresponding LRT-based detectors where many new detectors are derived and have not been explored in the literature. Table VI lists these detectors according to taxonomy based on the four types of data preprocessing, i.e., DS, covariance matrix-based data whitening via $\mathbf{K}^{-1/2}$, correlation matrix-based data whitening via $\mathbf{R}^{-1/2}$ and OSP where those detectors highlighted by “*” are new detectors which have never been reported in the literature.

- 1) Using DS to produce DS-space: *NDS-SNR = NAMD, *GDS-SNR, *NGDS-SNR = (NAMD)², *DS-SA².
- 2) Using $\mathbf{K}^{-1/2}$ -whitening matrix to produce K-space: *NK-SNR = NLRT, *GK-SNR = GLRT = AMF, *NGK-SNR = ASD = *(NLRT)², NMF = *K-SA, ACE = *K-SA².

TABLE V
LIST OF SA-DERIVED DETECTORS

SA-based detection methods	Subpixel/Mixed pixel detectors	Equation
K-SA=NMF	$\delta^{\text{NMF}}(\mathbf{r}) = \frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1/2} (\mathbf{r}^T \mathbf{K}^{-1} \mathbf{r})^{-1/2}} = \cos \theta = \delta^{\text{K-SA}}(\mathbf{r})$	(81)
K-SA ² =ACE	$\delta^{\text{ACE}}(\mathbf{r}) = \frac{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2}{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})(\mathbf{r}^T \mathbf{K}^{-1} \mathbf{r})} = \cos^2 \theta = \delta^{\text{K-SA}^2}(\mathbf{r})$	(82)
DS-SA ²	$\delta^{\text{K-SA}^2}(\mathbf{r}) = \frac{((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b))^2}{((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b))((\mathbf{r} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b))} = \cos^2 \hat{\theta}$	(83)
R-SA ²	$\delta^{\text{R-SA}^2}(\mathbf{r}) = \frac{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r})^2}{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t})(\mathbf{r}^T \mathbf{R}^{-1} \mathbf{r})} = \cos^2 \bar{\theta}$	(85)

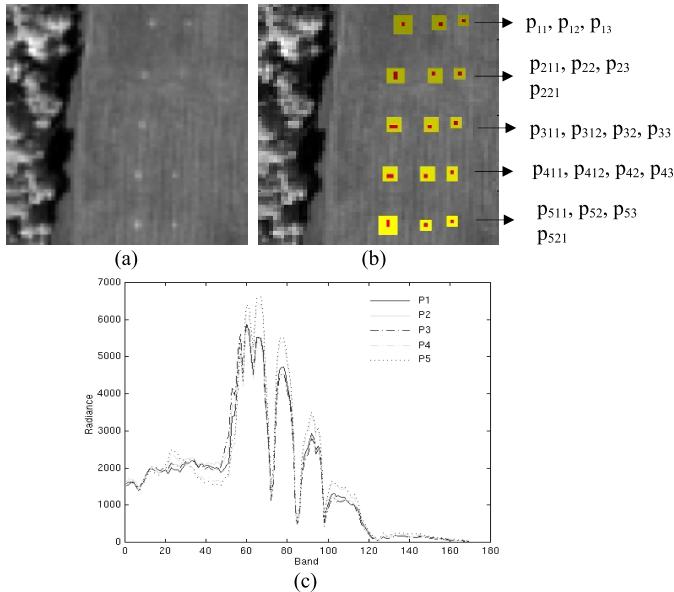


Fig. 1. (a) HYDICE panel scene which contains 15 panels. (b) Ground truth map of spatial locations of 19 R panel pixels. (c) Spectra of \mathbf{p}_1 , \mathbf{p}_2 , \mathbf{p}_3 , \mathbf{p}_4 , and \mathbf{p}_5 .

- 3) Using $\mathbf{R}^{-1/2}$ -whitening matrix to produce R space:
 $^* \text{NGR-SNR} = (\text{CEM})^2$, $^* \text{R-SA}^2$.
- 4) Using OSP for annihilation of undesired target signatures, \mathbf{U} : $^* \text{OSP-GLRT}$, $^* \text{NOSP-GK-SNR}$.

It should be noted that detection maps and ROC analysis are not affected by constants. So, the results of GDS-SNR and NGDS-SNR are the same with their difference by the constant $\kappa = ((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b))^{-1}$, so are GK-SNR = GLRT = AMF and NGK-SNR = (NLRT)² by the constant $\kappa = (\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})^{-1}$, and GR-SNR and NGR-SNR = (CEM)² by the constant, $\kappa = (\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t})^{-1}$. In this case, only one of them was included for experiments. As a result, a total of 12 detectors were implemented: 1)NAMD; 2) (NAMD)²; 3) NLRT; 4) ASD = (NLRT)²; 5) CEM; 6) (CEM)²; 7) OSP-GLRT; 8) NOSP-GK-SNR; 9) NMF; 10) ACE; 11) DS-SA²; and 12) R-SA². Nevertheless, a note is worthwhile. The

constant κ may not matter in detection but it does make a great difference and has significant impact on unmixing target pixels in terms of their detected abundance fractions.

A. Hyperspectral Digital Imagery Collection Experiment Data

An airborne hyperspectral digital imagery collection experiment (HYDICE) scene shown in Fig. 1(a) was used to analyze comparative performance among detectors discussed in Sections III–VII. It was collected in August 1995 from a flight altitude of 10 000 ft. here are 15 square panels in Fig. 1(a) with three different sizes, 3 m \times 3 m, 2 m \times 2 m, and 1 m \times 1 m, respectively. Owing to the ground sampling distance of approximately 1.56 m, each of panels in the first column except the first row contains two-panel pixels highlighted by red, p_{211} , p_{221} in row 2, p_{311} , p_{312} in row 3, p_{411} , p_{412} , p_{42} , p_{43} in row 4, p_{511} , p_{521} in row 5 as shown in Fig. 1. All the remaining 11 panels in Fig. 1 contain one single panel pixel for each panel also highlighted by red, p_{11} , p_{12} , p_{13} in row 1, p_{22} , p_{23} in row 2, p_{32} , p_{33} in row 3, p_{42} , p_{43} in row 4, p_{52} , p_{53} in row 5. Therefore, there are a total of 19 red (R) panel pixels. Fig. 1(b) shows their precise spatial locations with the pixels in yellow (Y pixels) indicating panel pixels mixed with the BKG. Fig. 1(c) plots the five-panel spectral signatures, \mathbf{p}_1 , \mathbf{p}_2 , \mathbf{p}_3 , \mathbf{p}_4 , and \mathbf{p}_5 obtained from Fig. 1(b), where the i th panel signature, denoted by \mathbf{p}_i was generated by averaging the red panel center pixels in row i . The best advantage of using this image scene is the ground truth provided by Fig. 1(b) about the precise spatial locations of each of 19 R panel pixels shown in Fig. 1(b) so that the exact spatial position of each panel pixel can be located to be used to evaluate each detector in the detection of 19 R panel pixels up to pixel level.

Fig. 2 shows the detection maps of panels in each row in Fig. 1(b) which were produced by detectors in Table VI using the desired target signature \mathbf{t} obtained from Fig. 1(c).

By visual inspection of detection maps in Fig. 2 it was easy to identify (a) NAMD, (e) CEM, (f) GR-SNR = (CEM)², (g) OSP-GLRT, (h) NOSP-GK-SNR, (k) DS-SA², (l) R-SA²

TABLE VI
TAXONOMY OF DETECTORS ACCORDING TO DATA FOUR TYPES OF PREPROCESSING, DS, K-WHITENING, R-WHITENING, AND OSP

Data pre-processing	detectors	Detector equation
SNR calculated in sphered data (DS-SNR)	*NDS-SNR=NAMD	$\delta^{\text{NAMD}}(\mathbf{r}) = \frac{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)}{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b)}$
	*GDS-SNR	$\delta^{\text{GDS-SNR}}(\mathbf{r}) = \frac{[(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)]^2}{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b)}$
	*NGDS-SNR=(NAMD) ²	$\delta^{\text{NGDS-SNR}}(\mathbf{r}) = \left(\frac{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b)}{(\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b)} \right)^2 = (\delta^{\text{NAMD}}(\mathbf{r}))^2$
SNR calculated in $\mathbf{K}^{-1/2}$ -whitened data (K-SNR)	*NK-SNR=NLRT	$\delta^{\text{NLRT}}(\mathbf{r}) = \frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}}$
	*GK-SNR =GLRT=AMF	$\delta^{\text{GLRT}}(\mathbf{r}) = \frac{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}}$
	*NGK-SNR = *(NLRT) ² =ASD	$\delta^{\text{ASD}}(\mathbf{r}) = \left(\frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t}} \right)^2$
SNR calculated in $\mathbf{R}^{-1/2}$ -whitened data R-SNR	*NR-SNR =CEM	$\delta^{\text{CEM}}(\mathbf{r}) = \frac{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}}$
	*GR-SNR	$\delta^{\text{GLRT}}(\mathbf{r}) = \frac{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r})^2}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}}$
	*NGR-SNR=(CEM) ²	$\delta^{\text{NGR-SNR}}(\mathbf{r}) = \left(\frac{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r}}{\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t}} \right)^2$
Data preprocessed by OSP OSP	*OSP-GLRT	$\delta^{\text{OSP-GLRT}}(\mathbf{r}) = \left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p \right)^{-1} \left(\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{P}_{\mathbf{U}_p}^T \mathbf{r} \right)^2 \text{ with } \tilde{\mathbf{K}} \text{ defied in (22)}$
	*NOSP-GK-SNR	$\delta^{\text{NOSP-GK-SNR}}(\mathbf{r}) = \left(\frac{\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{P}_{\mathbf{U}_p}^T (\mathbf{r} - \boldsymbol{\mu}_b)}{\mathbf{t}_p^T \tilde{\mathbf{K}}^{-1} \mathbf{t}_p} \right)^2 \text{ with } \tilde{\mathbf{K}} \text{ defied in (22)}$
SA calculated in sphered data	*DS-SA ²	$\delta^{\text{DS-SA}^2}(\mathbf{r}) = \frac{((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b))^2}{((\mathbf{t} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{t} - \boldsymbol{\mu}_b))((\mathbf{r} - \boldsymbol{\mu}_b)^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}_b))} = \cos^2 \hat{\theta}$
SA calculated in $\mathbf{K}^{-1/2}$ -whitened data	NMF=K-SA	$\delta^{\text{NMF}}(\mathbf{r}) = \delta^{\text{K-SA}}(\mathbf{r}) = \frac{\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r}}{\left(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t} \right)^{-1/2} \left(\mathbf{r}^T \mathbf{K}^{-1} \mathbf{r} \right)^{-1/2}} = \cos \theta$
	ACE=K-SA ²	$\delta^{\text{NMF}}(\mathbf{r}) = \frac{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{r})^2}{(\mathbf{t}^T \mathbf{K}^{-1} \mathbf{t})(\mathbf{r}^T \mathbf{K}^{-1} \mathbf{r})} = \cos^2 \theta$
SA calculated in $\mathbf{R}^{-1/2}$ -whitened data	*R- SA ²	$\delta^{\text{R-SA}^2}(\mathbf{r}) = \frac{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{r})^2}{(\mathbf{t}^T \mathbf{R}^{-1} \mathbf{t})(\mathbf{r}^T \mathbf{R}^{-1} \mathbf{r})} = \cos^2 \bar{\theta}$

were among the best detectors, while (a) NAMD, (e) CEM, (h) NOSP-GK-SNR did not suppress BKG as well as the other best detectors did. This is because when the latter detected panel pixels in row 2, they also detected panel pixels in

row 3 and vice versa. The worst cases were produced by NMF and ACE with (c) NLRT and (d) ASD right in between where the BKG effect had significant affected their detection performance.

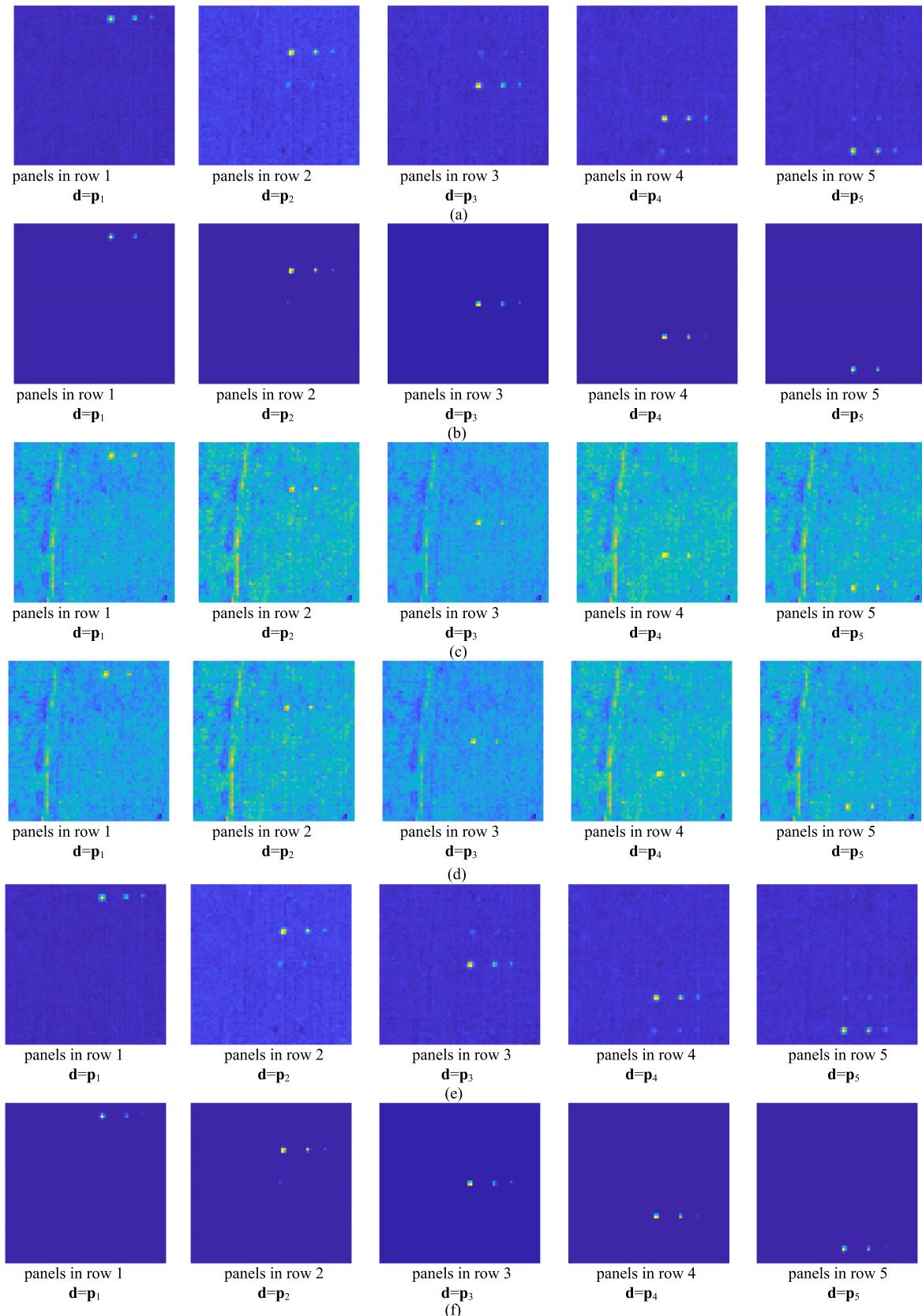


Fig. 2. Detection maps of 19 panel pixels in five rows by NAMD, $(\text{NAMD})^2$, NLRT, ASD = $(\text{NLRT})^2$, CEM, NGR-SNR = $(\text{CEM})^2$, OSP-GLRT, NOSP-GK-SNR, NMF = K-SA, ACE = $(\text{NMF})^2$, DS-SA 2 , R-SA 2 . (a) Detection maps of NAMD. (b) Detection maps of $(\text{NAMD})^2$. (c) Detection maps of NLRT. (d) Detection maps of ASD = $(\text{NLRT})^2$. (e) Detection maps of CEM. (f) Detection maps of NGR-SNR = $(\text{CEM})^2$. (g) Detection maps of OSP-GLRT. (h) Detection maps of NOSP-GK-SNR. (i) Detection maps of NMF. (j) Detection maps of ACE = $(\text{NMF})^2$. (k) Detection maps of DS-SA 2 . (l) Detection maps of R-SA 2 .

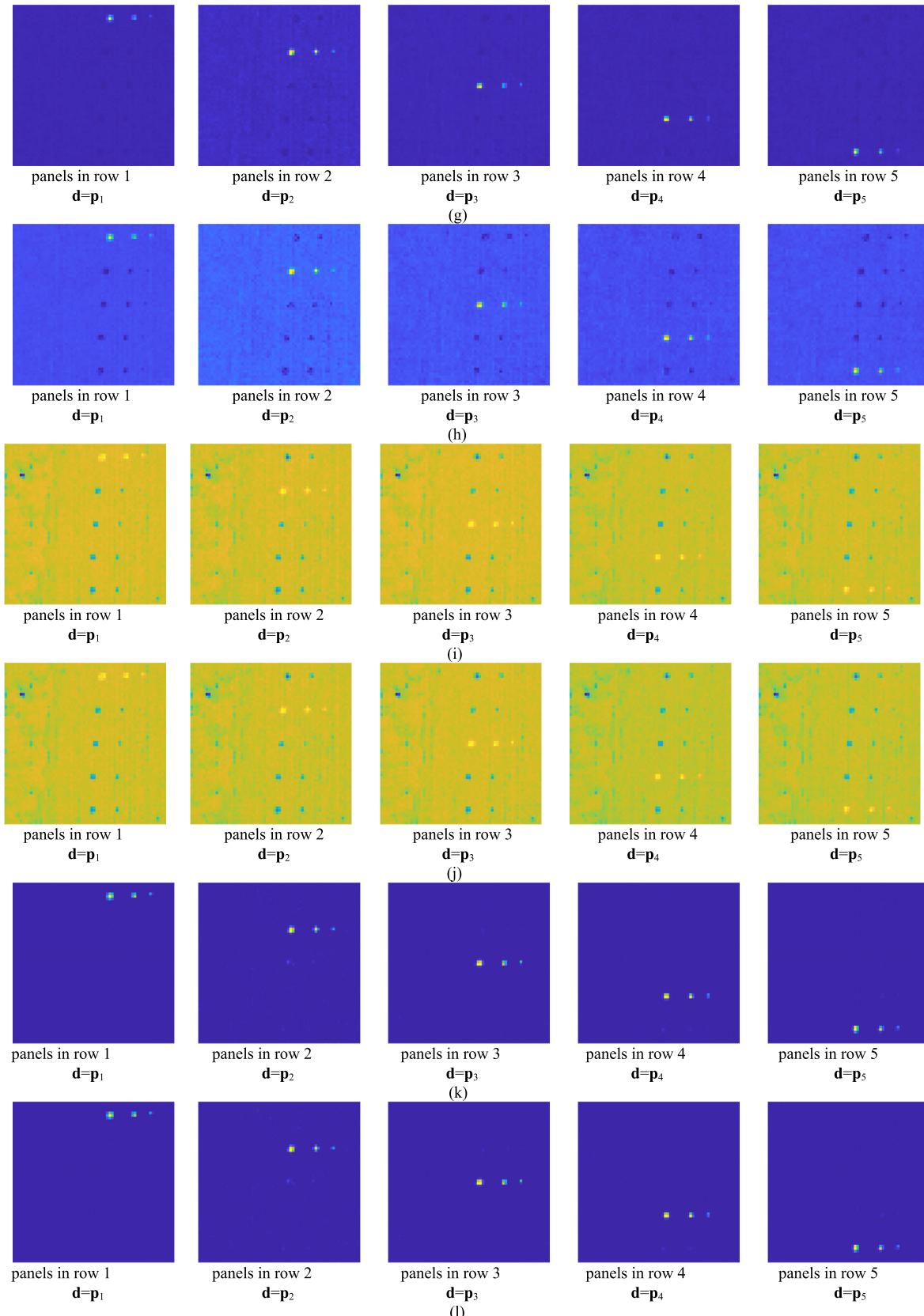


Fig. 2. (Continued.) Detection maps of 19 panel pixels in five rows by NAMD, $(\text{NAMD})^2$, NLRT, $\text{ASD} = (\text{NLRT})^2$, CEM, $\text{NGR-SNR} = (\text{CEM})^2$, OSP-GLRT, NOSP-GK-SNR, NMF = K-SA, ACE = $(\text{NMF})^2$, DS-SA 2 , R-SA 2 . (a) Detection maps of NAMD. (b) Detection maps of $(\text{NAMD})^2$. (c) Detection maps of NLRT. (d) Detection maps of $\text{ASD} = (\text{NLRT})^2$. (e) Detection maps of CEM. (f) Detection maps of $\text{NGR-SNR} = (\text{CEM})^2$. (g) Detection maps of OSP-GLRT. (h) Detection maps of NOSP-GK-SNR. (i) Detection maps of NMF. (j) Detection maps of ACE = $(\text{NMF})^2$. (k) Detection maps of DS-SA 2 . (l) Detection maps of R-SA 2 .

TABLE VII

AUC VALUES CALCULATED FROM THE THREE 2-D ROC CURVES FOR FIVE ROWS DETECTION RESULTS OBTAINED FOR HYDICE DATA FROM FIG. 2

d	detector	AUC_(D,F)	AUC_(D,G)	AUC_(F,G)	AUC_{TD}	AUC_{BS}	AUC_{TDBS}	AUC_{SNPR}	AUC_{ODP}
p₁	(a)NAMD	0.9993	0.5947	0.0085	1.5940	0.9904	0.5863	70.2083	1.5855
	(b)(NAMD) ²	0.9993	0.4526	0.0005	1.4519	0.9984	0.4521	885.5845	1.4514
	(c)NLRT	0.9950	0.7750	0.3791	1.7701	0.6135	0.3960	2.0447	1.3910
	(d)ASD=(NLRT) ²	0.9950	0.7644	0.3617	1.7594	0.6309	0.4027	2.1133	1.3977
	CEM	0.9993	0.5953	0.0084	1.5946	0.9905	0.5869	70.8589	1.5862
	NGR-SNR=(CEM) ²	0.9993	0.4530	0.0005	1.4523	0.9984	0.4525	886.8226	1.4518
	(g)OSP-GLRT	0.9993	0.4947	0.0179	1.4940	0.9810	0.4768	27.6108	1.4761
	(h)NOSP-GK-SNR	0.9993	0.6474	0.1289	1.6466	0.8700	0.5185	5.0228	1.5177
	(i)NMF=K-SA	0.9992	0.9484	0.7674	1.9476	0.2314	0.1810	1.2359	1.1802
	(j)ACE=(NMF) ²	0.9992	0.9451	0.7554	1.9443	0.2434	0.1898	1.2512	1.1890
	(k)DS-SA ²	0.9994	0.7416	0.0019	1.7411	0.9972	0.7397	385.6793	1.7392
	(l)R-SA ²	0.9994	0.7426	0.0019	1.7420	0.9972	0.7407	388.7317	1.7401
p₂	(a)NAMD	0.9998	0.8310	0.0169	1.8309	0.9828	0.8141	49.0473	1.8139
	(b)(NAMD) ²	0.9998	0.7599	0.0010	1.7598	0.9988	0.7590	791.7636	1.7588
	(c)NLRT	0.9950	0.7818	0.4320	1.7768	0.5613	0.3498	1.8098	1.3448
	(d)ASD=(NLRT) ²	0.9950	0.7716	0.4160	1.7666	0.5773	0.3555	1.8546	1.3505
	CEM	0.9998	0.8325	0.0169	1.8323	0.9828	0.8156	49.2400	1.8154
	NGR-SNR=(CEM) ²	0.9998	0.7620	0.0010	1.7618	0.9988	0.7610	789.3837	1.7609
	(g)OSP-GLRT	0.9998	0.7982	0.0391	1.7980	0.9606	0.7590	20.4062	1.7589
	(h)NOSP-GK-SNR	0.9998	0.8675	0.1935	1.8673	0.8062	0.6740	4.4821	1.6738
	(i)NMF=K-SA	0.9999	0.9755	0.7623	1.9754	0.2375	0.2132	1.2796	1.2131
	(j)ACE=(NMF) ²	0.9999	0.9740	0.7503	1.9739	0.2496	0.2237	1.2982	1.2236
	(k)DS-SA ²	0.9998	0.8135	0.0027	1.8133	0.9970	0.8108	299.2357	1.8106
	(l)R-SA ²	0.9998	0.8138	0.0027	1.8137	0.9971	0.8111	300.8682	1.8110
p₃	(a)NAMD	0.9997	0.7638	0.0108	1.7635	0.9888	0.7530	70.4848	1.7527
	(b)(NAMD) ²	0.9997	0.6338	0.0006	1.6335	0.9990	0.6331	1024.5582	1.6328
	(c)NLRT	0.9966	0.8045	0.3568	1.8011	0.6387	0.4477	2.2547	1.4443
	(d)ASD=(NLRT) ²	0.9966	0.7942	0.3375	1.7908	0.6580	0.4567	2.3533	1.4534
	CEM	0.9997	0.7658	0.0108	1.7655	0.9888	0.7550	70.8975	1.7547
	NGR-SNR=(CEM) ²	0.9997	0.6361	0.0006	1.6358	0.9990	0.6355	1022.9401	1.6352
	(g)OSP-GLRT	0.9998	0.6773	0.0230	1.6770	0.9766	0.6542	29.3938	1.6540
	(h)NOSP-GK-SNR	0.9998	0.8004	0.1498	1.8002	0.8499	0.6507	5.3448	1.6504
	(i)NMF=K-SA	0.9999	0.9918	0.7700	1.9917	0.2299	0.2218	1.2881	1.2217
	(j)ACE=(NMF) ²	0.9999	0.9912	0.7578	1.9912	0.2421	0.2334	1.3080	1.2333
	(k)DS-SA ²	0.9998	0.8539	0.0019	1.8537	0.9979	0.8520	448.7642	1.8518
	(l)R-SA ²	0.9998	0.8554	0.0019	1.8553	0.9979	0.8535	450.5786	1.8534
p₄	(a)NAMD	0.9998	0.8149	0.0105	1.8147	0.9892	0.8044	77.3337	1.8042
	(b)(NAMD) ²	0.9998	0.7286	0.0005	1.7284	0.9992	0.7281	1328.3001	1.7279
	(c)NLRT	0.9729	0.8526	0.4520	1.8255	0.5119	0.4006	1.8862	1.3735
	(d)ASD=(NLRT) ²	0.9729	0.8458	0.4342	1.8188	0.5297	0.4116	1.9478	1.3845
	CEM	0.9998	0.8149	0.0104	1.8147	0.9894	0.8045	78.5316	1.8044
	NGR-SNR=(CEM) ²	0.9998	0.7285	0.0005	1.7283	0.9992	0.7279	1329.8113	1.7278
	(g)OSP-GLRT	0.9998	0.7501	0.0179	1.7499	0.9818	0.7322	41.8680	1.7320
	(h)NOSP-GK-SNR	0.9998	0.8375	0.1318	1.8373	0.8680	0.7057	6.3560	1.7056
	(i)NMF=K-SA	0.9998	0.9604	0.7482	1.9602	0.2515	0.2122	1.2836	1.2120
	(j)ACE=(NMF) ²	0.9998	0.9579	0.7350	1.9577	0.2647	0.2228	1.3032	1.2227
	(k)DS-SA ²	0.9998	0.8444	0.0018	1.8442	0.9980	0.8426	478.3755	1.8425
	(l)R-SA ²	0.9998	0.8448	0.0017	1.8446	0.9980	0.8431	484.2071	1.8429
p₅	(a)NAMD	0.9998	0.7372	0.0104	1.7369	0.9893	0.7268	71.0737	1.7266
	(b)(NAMD) ²	0.9998	0.6213	0.0005	1.6210	0.9992	0.6208	1312.0059	1.6206
	(c)NLRT	0.9974	0.8635	0.4157	1.8610	0.5808	0.4478	2.0772	1.4452
	(d)ASD=(NLRT) ²	0.9974	0.8566	0.3987	1.8540	0.5979	0.4579	2.1485	1.4553
	CEM	0.9998	0.7376	0.0103	1.7373	0.9894	0.7273	71.4553	1.7270
	NGR-SNR=(CEM) ²	0.9998	0.6216	0.0005	1.6214	0.9992	0.6212	1315.3577	1.6209
	(g)OSP-GLRT	0.9998	0.6490	0.0197	1.6488	0.9800	0.6293	32.9322	1.6291
	(h)NOSP-GK-SNR	0.9998	0.7673	0.1370	1.7670	0.8626	0.6302	5.5988	1.6300
	(i)NMF=K-SA	0.9998	0.9555	0.7525	1.9553	0.2472	0.2030	1.2698	1.2028
	(j)ACE=(NMF) ²	0.9998	0.9528	0.7398	1.9526	0.2600	0.2130	1.2879	1.2128
	(k)DS-SA ²	0.9998	0.7742	0.0018	1.7739	0.9979	0.7724	434.8178	1.7722
	(l)R-SA ²	0.9998	0.7747	0.0018	1.7744	0.9979	0.7729	438.8971	1.7727

The visual inspection-based qualitative analysis from Fig. 2 could not differentiate among the best detectors. The detection measures introduced in Section VIII are particularly designed for this purpose of quantitative analysis. According to [9], we first produced 3-D ROC curves of the detection maps in Fig. 2 and then calculated their corresponding $AUC_{(D,F)}$ values of 2-D ROC curves of (P_D, P_F) , $AUC_{(D,\tau)}$ values of 2-D ROC curves of (P_D, τ) , and $AUC_{(F,\tau)}$ values of 2-D ROC curves of (P_F, τ) . Using these three AUC values we further calculated the five detection measures based on (87)–(91). Table VII tabulates all these AUC results where the best results are boldfaced. Interestingly, all test detectors produced nearly the same $AUC_{(D,F)}$ values approximately 0.99 with their differences less than 10^{-2} including the worst detectors, the two worst detectors, (i) NMF and (j) ACE and three other not best detectors, (a) NLRT, (i) GLRT, and (j) ASD. However, if we factored their $AUC_{(F,\tau)}$ values in evaluating their BS, (i) NMF and (j) ACE produced nearly 0.73–0.77 false alarm probabilities which were almost twice much 0.36–0.45 produced by (a) NLRT, (i) GLRT, and (j) ASD. By contrast, the best detector was GR-SNR produced only 0.0005–0.0006 with three other best detectors, (b) NAMD, (k) DS-SA², and (l) R-SA² which also produced only 0.0018–0.0027. On the other hand. These two detectors (i) NMF and (j) ACE also produced the highest $AUC_{(D,\tau)}$ values of 2-D ROC curves of (P_D, τ) greater than 0.95 and highest AUC_{TD} values regardless of their also highest $AUC_{(F,\tau)}$ values of 2-D ROC curves of (P_F, τ) around 0.73–0.77. These experiments demonstrated that the common practice in the HTD literature using only $AUC_{(D,F)}$ of 2-D ROC curves of (P_D, P_F) as an evaluation measure is not reliable and may also mislead conclusion since a very high value of $AUC_{(D,F)}$ produces a very high $AUC_{(D,\tau)}$ of P_D but also a very high-value $AUC_{(F,\tau)}$ of P_F as shown by NMF and ACE in Table VII, both of which produced the highest values of $AUC_{(D,\tau)}$ and $AUC_{(F,\tau)}$ but turned out to be the worst detectors. This mainly results from (4) and (5) which are calculated by the same threshold τ . Details on 3-D ROC analysis can be found in [9]. If we further examined the values of AUC_{BS} , AUC_{TDBS} , and AUC_{ODP} , (i) NMF and (j) ACE were the worst and (b) NAMD, (h) NGR-SNR, (k) DS-SA², (l) R-SA² were among the best. Specifically, (k) DS-SA², (l) R-SA² produced the highest AUC_{SNPR} values way above in an exponential order than that produced by other detectors. This indicated that the $AUC_{(F,\tau)}$ and AUC_{SNPR} values are only two effective measures that could validate the qualitative performance by visual inspection in Fig. 2. By all things considered based on Table VI, we could conclude that the best nonlinear matched detector would be (k) DS-SA², (l) R-SA². However, as for best linear MFs, the best ones would be (b) NAMD, (h) NGR-SNR, and (m) CEM.

B. San Diego Airport Data

A second data set used for experiments was an airborne visible/infrared imaging spectrometer (AVIRIS) acquired scene, which is a San Diego airport scene in CA, USA with pseudo-color shown in Fig. 3(a).

It has a size of 400×400 pixels with a 3.5 m spatial resolution and 224 spectral channels in wavelengths ranging from 370 to 2510 nm. After removing the bands that correspond to

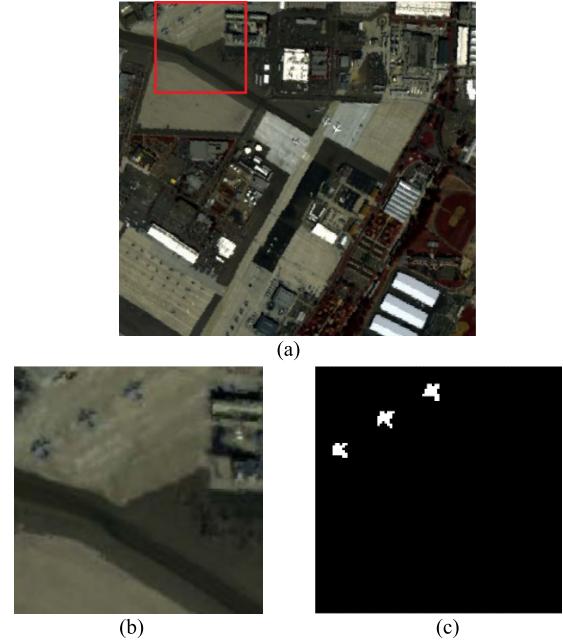


Fig. 3. San Diego airport scene. (a) Pseudo-color image of the whole scene. (b) Pseudo-color image of the selected area. (c) Ground truth map.

the water absorption regions, low SNR, and bad bands (1–6, 33–35, 97, 107–113, 153–166, and 221–224), 189 available bands of the data were retained in the experiments. An area of 100×100 pixels at the upper left corner of the scene was selected as the test image shown in Fig. 3(b) along with its ground-truth map shown in Fig. 3(c). It is an urban scene in which the main BKG materials are roof, shadow, and grass. There are three airplanes in the image, which consist of 58 pixels and account for 0.33% of the test image as shown in Fig. 3(b). The spectral signatures of these planes are different from those of the main BKG as shown in Fig. 3(c), and these pixels were, therefore, selected as targets to be detected. This scene is less interesting than the HYDICE 15-panel data in Fig. 1 because the only ground truth is available is three airplanes which are large targets.

Since there are no undesired targets in \mathbf{U} like HYDICE data which can be used for experiments, OSP-GLRT and NOSP-GSNR were not implemented for this scene. Following the same experiments conducted for HYDICE data, Fig. 4 shows the detection maps of three airplanes in Fig. 4 produced by ten detectors using the desired target signature \mathbf{t} obtained from the airplanes in Fig. 3(c). By visual inspection of detection results of three airplanes in Fig. 4(a)–(j) where the three best detectors, NAMD, (NAMD)², NLRT, ASD = (NLRT)², CEM, NGR-SNR = (CEM)², NMF, ACE = (NMF)², DS-SA², and R-SA².

According to Fig. 4 we could divide all the ten detectors into three categories, good, fair, and poor by visual inspection for qualitative analysis, but could not differentiate the detectors in each category. Therefore, once again, Table VIII tabulates all these AUC results obtained from the detection maps in Fig. 4 where the best results are boldfaced. Like HYDICE data experiments, NMF produced the highest $AUC_{(D,F)}$ and the worst BS with the highest $AUC_{(F,\tau)}$ but lowest AUC_{BS} , AUC_{TDBS} , and AUC_{ODP} and lowest AUC_{SNPR} . On the other hand, (e) (NAMD)² produced the lowest $AUC_{(F,\tau)}$ values and highest AUC_{SNPR} values, which implied that (NAMD)²

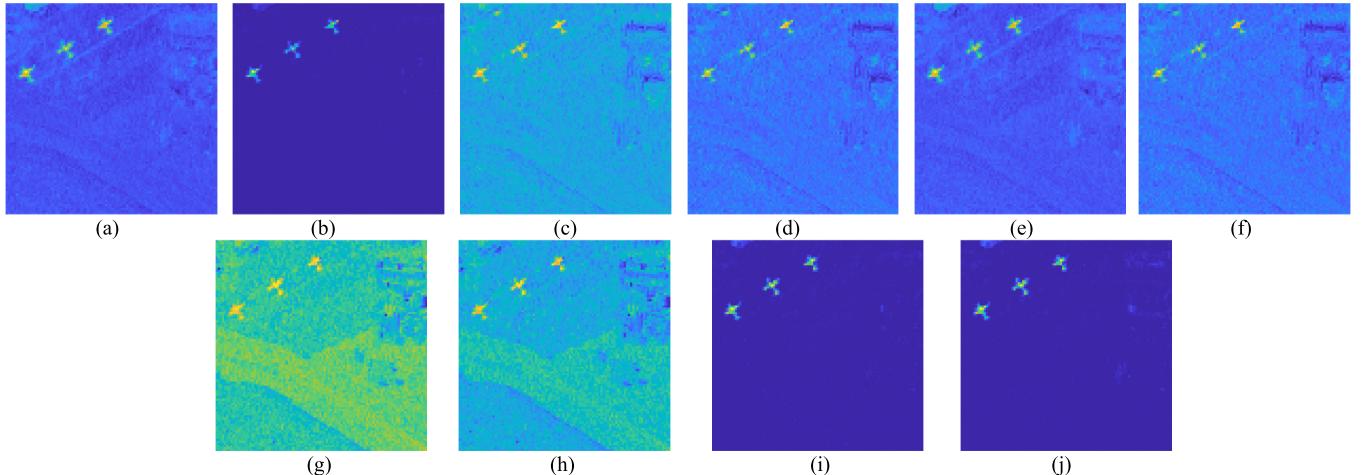


Fig. 4. Detection maps of San Diego airport data. (a) NAMD. (b) $(\text{NAMD})^2$. (c) NLRT. (d) $\text{ASD} = (\text{NLRT})^2$. (e) CEM. (f) $\text{NGR-SNR} = (\text{CEM})^2$. (g) $\text{NMF} = \text{K-SA}$. (h) $\text{ACE} = (\text{NMF})^2$. (i) DS-SA^2 . (j) R-SA^2 .

TABLE VIII
AUC VALUES CALCULATED FROM THE THREE 2-D ROC CURVES FOR FIVE ROWS DETECTION
RESULTS OBTAINED FOR AN DIEGO AIRPORT FROM FIG. 5

detector	$\text{AUC}_{(\text{D},\text{F})}$	$\text{AUC}_{(\text{D},\tau)}$	$\text{AUC}_{(\text{F},\tau)}$	AUC_{TD}	AUC_{BS}	AUC_{TDBS}	AUC_{SNPR}	AUC_{ODP}
(a)NAMD	0.9766	0.4768	0.0295	1.4533	0.9468	0.4473	16.1793	1.4239
(b) $(\text{NAMD})^2$	0.9766	0.2886	0.0019	1.2652	0.9744	0.2867	154.7424	1.2633
(c)NLRT	0.9135	0.6576	0.3775	1.5711	0.5349	0.2801	1.7418	1.1936
(d)ASD= $(\text{NLRT})^2$	0.9135	0.5366	0.2286	1.4500	0.6838	0.3079	2.3468	1.2214
(e)CEM	0.9901	0.4862	0.0301	1.4763	0.9599	0.4561	16.1598	1.4462
(f)NGR-SNR= $(\text{CEM})^2$	0.9901	0.2929	0.0020	1.2830	0.9880	0.2909	143.7513	1.2810
(g)NMF=K-SA	0.8958	0.7774	0.5625	1.6732	0.3321	0.2148	1.3819	1.1107
(h)ACE= $(\text{NMF})^2$	0.8958	0.6961	0.4344	1.5920	0.4602	0.2617	1.6025	1.1576
(i)DS-SA ²	0.9772	0.4032	0.0038	1.3803	0.9731	0.3994	106.8681	1.3766
(j)R-SA ²	0.9911	0.4087	0.0041	1.3998	0.9868	0.4046	98.7384	1.3956

performed the best BS. Interestingly, the three detectors (e) $(\text{NAMD})^2$, (i) DS-SA^2 , and (j) R-SA^2 which produced the best qualitative analysis by visual inspection did not produce very high AUC values in terms of (D,τ) , TD, TDBS, and ODP due to their poor TD even they produced very high $\text{AUC}_{(\text{D},\text{F})}$ values. These experiments demonstrated that BS and SNPR are the only two effective detection measures to be able to evaluate the degree of BS. Specifically, their AUC_{SNPR} values were way above that produced by other detectors ten times NAMD and CEM but 100 times other detectors. Nevertheless, overall speaking, among the best detectors are (e) CEM, (f) $\text{NGR-SNR} = (\text{CEM})^2$, (i) DS-SA^2 , and (j) R-SA^2 . Since (i) DS-SA^2 and (j) R-SA^2 are nonlinear MFs, the best linear filter would be (k) CEM according to AUC_{ODP} .

C. Discussions

The experiments performed in Sections VIII-A and VIII-B provide interesting observations and findings which can be analyzed from three aspects.

1) *BKG Effects*: Owing to significantly improved spectral resolution, a hyperspectral imaging sensor is capable of uncovering many subtle substances which cannot be identified by prior knowledge or visual inspection. In statistical detection theory, signal detection in noise is generally formulated by

a binary hypothesis testing problem which assumes BKG to be an additive Gaussian noise. However, a hyperspectral image generally has a very complicated BKG including many unknown and unidentified interferers other than noise. As a result, one of the major challenging issues arising in HTD is how to cope with BKG effects.

A general approach is to assume that BKG can be characterized by the second-order data statistics so that removing BKG effects can be accomplished by decorrelating second-order data statistics. Three methods are proposed in this article for doing so.

- 1) DS which removes the first two orders of data statistic by (27).
- 2) Data whitening using $\mathbf{K}^{-1/2}$ by (37).
- 3) Data whitening using $\mathbf{R}^{-1/2}$ by (43).

All of these three methods have been used to derive LRT, SNR, and SA detectors. The results were shown to be very impressive.

2) *Effectiveness of a Detector*: The AUC of a 2-D ROC curve of $(P_{\text{D}}, P_{\text{F}})$, $\text{AUC}_{(\text{D},\text{F})}$ has been widely used to evaluate the detection performance of a detector. Unfortunately, it is not reliable in [9]. This can be easily seen from the experimental results in Tables VII and VIII.

- 1) For HYDICE in Table VII, $AUC_{(D,F)}$ values were very close and nearly 0.99 with their differences by 0.005, in which case it is very difficult to determine which one is the best detector.
- 2) For San Diego airport data in Table VIII, $R-SA^2$, CEM, $NGR-SNR = (CEM)^2$ produced the highest $AUC_{(D,F)}$ values followed by NAMD, $(NAMD)^2$, and DS-SA² with $AUC_{(D,F)}$ values around 0.97 and the rest with $AUC_{(D,F)}$ values 0.89–0.91.

3) *TD and BS*: In addition to $AUC_{(D,F)}$ discussed above, there are also five other detection measures, $AUC_{(D,\tau)}$ and AUC_{TD} , $AUC_{(F,\tau)}$, AUC_{BS} , and AUC_{SNPR} , which can be also used for performance evaluation.

- 1) TD is determined by $AUC_{(D,\tau)}$ and AUC_{TD} in which case, NMF and ACE = $(NMF)^2$ were the best for both data sets.
- 2) BS is determined by $AUC_{(F,\tau)}$, AUC_{BS} , and AUC_{SNPR} , in which case, NAMD and $NGR-SNR = (CEM)^2$ were the best for both data sets.
- 3) Overall performance is determined by AUC_{TDBS} and AUC_{ODP} . For HYDICE, $R-SA^2$, DS-SA² were the best and performed slightly better than CEM and NAMD. But for San Diego airport, CEM and NAMD were the best and performed slightly better than $R-SA^2$, DS-SA².

Therefore, as discussed above, it is very obvious that solely depending on one single detection measure to evaluate a detector is biased and is very likely to mislead conclusions. Also, it seemed that the visual inspection from detection maps in Figs. 2 and 4 could be only used to conform to the BKG suppressibility of a detector but not its TD, TDBS, and ODP. In the end, no detector can claim to perform optimally for all detection measures.

Finally, according to the experiments, OSP-GLRT and OSP-GK-SNR performed reasonably well but did not perform as well as the best detectors did. This was mainly because the undesired target signatures used in \mathbf{U} did not faithfully represent the BKG for annihilation. It is believed that if \mathbf{U} is appropriately determined, the detection performance of OSP-based GLRT and GSNR detectors will be significantly improved. A recent work using the low-rank and sparse matrix decomposition (LRaSMD) model to address this issue of \mathbf{U} was reported in [36]. Future work along this line is worth being further investigated.

X. CONCLUSION

This article revisits the hypothesis testing-based LRT detection for HTD and also rederives its theory, specifically, subpixel target and mixed-pixel target detection. However, such an approach requires prior knowledge of probability distributions which are generally assumed to be Gaussian distributed. Unfortunately, in practical applications obtaining prior knowledge and making Gaussian assumptions are not realistic. To address this issue two new HTD theories are also derived, SNR-based detection theory which only requires the first two order data statistics without appealing for probability distributions, and SA-based detection theory, which does not require any assumption. To effectively utilize SNR-based and SA-based theories, three data whitening methods are also

particularly developed as preprocessing techniques to accomplish what the covariance matrix does in Gaussian distribution used by LRT detectors. Specifically, it has shown that Gaussian-based LRT can be interpreted under the umbrella of the SNR-based theory. This article provides a comprehensive treatment of these three detection theories including their detailed theoretical proofs and derivations as well as their comparative analysis.

Several contributions are summarized as follows.

- 1) It includes a very comprehensive theory and analysis for target detection. It offers a unified view of many target detectors developed in the literature.
- 2) It rederives theories for LRT/GLRT-based target detection using the very basic detection theory in [4] by formulating binary simple and composite hypothesis testing problems for full pure target detection and sub-pixel/mixed pixel target detection, respectively.
- 3) It derives theories for the SNR-based target detection and also shows that LRT/GLRT-based target detection can be unified in the context of the SNR/GSNR-based target detection theory. Specifically, the GSNR-based target detector can be shown to be identical to CEM.
- 4) It derives SA-based detectors as an alternative to MF-type of LRT-based and SNR-based detectors. In particular, the well-known ACE in [12], [16] can be shown to be an SA-detector.
- 5) It conducts a comprehensive study and analysis theoretically and experimentally among most target detectors currently being available and reported in the literature for comparative performance analysis.
- 6) The LRT-based, SNR-based and SA-based detection theories derived in this article provide fundamental concepts and insights in designing target detectors. These theories can be generalized in various directions, such as a) kernel extension; b) sparse representation; and c) spectral unmixing, each of which will offer more insights into the concept of the developed SNR-based theory. Specifically, investigations of extending the present work to anomaly detection, LRaSMD models are currently being undertaken.
- 7) Finally, since many existing target detectors currently being used in the literature have their roots in these three derived detection theories, this article gives a comprehensive and also tutorial review for comparative study and analysis on HTD.

ACKNOWLEDGMENT

The author would like to thank Mr. Jie Chen with UMBC for his help in implementing experiments presented in this article. In addition, the author would also like to thank anonymous reviewers for providing valuable suggestions to improve the article presentation.

REFERENCES

- [1] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. New York, NY, USA: Kluwer Academic, 2003.
- [2] C.-I Chang, *Hyperspectral Data Processing: Algorithm Design and Analysis*. Hoboken, NJ, USA: Wiley, 2013, ch. 2.

- [3] C.-I Chang, *Real-Time Progressive Hyperspectral Image Processing: Endmember Finding and Anomaly Detection*. New York, NY, USA: Springer, 2016.
- [4] H. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, USA: Springer-Verlag, 1991.
- [5] B. Picinbono, "On deflection as a performance criterion in detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 3, pp. 1072–1081, Jul. 1995.
- [6] J. C. Harsanyi and C.-I Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, Jul. 1994.
- [7] J. J. Settle, "On the relationship between spectral unmixing and subspace projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 4, pp. 1045–1046, Jul. 1996.
- [8] C.-I Chang, "Further results on relationship between spectral unmixing and subspace projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 3, pp. 1030–1032, May 1998.
- [9] C.-I Chang, "An effective evaluation tool for hyperspectral target detection: 3D receiver operating characteristic curve analysis," *IEEE Trans. Geosci. Remote Sens.*, early access, Sep. 25, 2020, doi: [10.1109/TGRS.2020.3021671](https://doi.org/10.1109/TGRS.2020.3021671).
- [10] J. C. Harsanyi, "Detection and classification of subpixel spectral signatures in hyperspectral image sequences," Ph.D. dissertation, Dept. Elect. Eng., Univ. Maryland, Baltimore County, MD, USA, 1993.
- [11] W. Farrand, "Mapping the distribution of mine tailings in the Coeur d'Alene River Valley, Idaho, through the use of a constrained energy minimization technique," *Remote Sens. Environ.*, vol. 59, no. 1, pp. 64–76, Jan. 1997.
- [12] D. Manolakis, M. Pieper, E. Truslow, T. Cooley, M. Brueggeman, and S. Lipson, "The remarkable success of adaptive cosine estimator in hyperspectral target detection," *Proc. SPIE*, vol. 8743, Art. no. 874302, May 2013, doi: [10.1117/12.2015392](https://doi.org/10.1117/12.2015392).
- [13] C.-I Chang, "A statistical detection theory approach to hyperspectral image classification: Performance analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2057–2074, Apr. 2019.
- [14] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–218, Jan. 1992.
- [15] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.
- [16] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2538–2541, Sep. 1999.
- [17] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3005–3014, Jan. 2001.
- [18] D. G. Manolakis, G. A. Shaw, and N. Keshava, "Comparative analysis of hyperspectral adaptive matched filter detectors," *Proc. SPIE*, vol. 4049, pp. 2–17, Aug. 2000, doi: [10.1117/12.41032](https://doi.org/10.1117/12.41032).
- [19] D. Manolakis, C. Siracusa, and G. Shaw, "Hyperspectral subpixel target detection using the linear mixing model," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 7, pp. 1392–1409, Jul. 2001.
- [20] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.
- [21] D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 79–116, 2003.
- [22] D. Manolakis, R. Lockwood, T. Cooley, and J. Jacobson, "Is there a best hyperspectral detection algorithm?" *Proc SPIE*, vol. 7334, Apr. 2009, Art. no. 733402.
- [23] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, and M. Brueggeman, "Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 24–33, Jan. 2014.
- [24] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.
- [25] Z. Zou and Z. Shi, "Hierarchical suppression method for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 330–342, Jan. 2016.
- [26] R. Zhao, Z. Shi, Z. Zou, and Z. Zhang, "Ensemble-based cascaded constrained energy minimization for hyperspectral target detection," *Remote Sens.*, vol. 11, no. 11, p. 1310, 2019, doi: [10.3390/rs1111310](https://doi.org/10.3390/rs1111310).
- [27] G. Camps-Valls, "Kernel spectral angle mapper," *Electron. Lett.*, vol. 52, no. 14, pp. 1218–1220, Jul. 2016.
- [28] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 629–640, Jun. 2011.
- [29] T. Guo, F. Luo, L. Zhang, B. Zhang, X. Tan, and X. Zhou, "Learning structurally incoherent background and target dictionaries for hyperspectral target detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3521–3533, 2020.
- [30] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Hoboken, NJ, USA: Wiley, 2001.
- [31] J. Wang and C.-I Chang, "Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1586–1600, Jun. 2006.
- [32] C. Yu, L. C.-Lee, C.-I Chang, B. Xue, M. Song, and J. Chen, "Band-specified virtual dimensionality: An orthogonal subspace approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2822–2832, May 2018.
- [33] T. M. Tu, C.-H. Chen, and C.-I Chang, "A posteriori least squares orthogonal subspace projection approach to weak signature extraction and detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 127–139, Jan. 1997.
- [34] C.-I Chang, "Chapter 3: Information-processed matched filters for hyperspectral target detection and classification," in *Hyperspectral Data Exploitation: Theory and Applications*. Hoboken, NJ, USA: Wiley, 2007, pp. 47–74.
- [35] 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 Sens.*, vol. 36, no. 3, pp. 898–912, May 1998.
- [36] C.-I Chang and J. Chen, "OSP using data spherling and low-rank and sparse matrix decomposition for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, early access, Feb. 2, 2021, doi: [10.1109/TGRS.2021.3053201](https://doi.org/10.1109/TGRS.2021.3053201).



Chein-I Chang (Life Fellow, IEEE) received the B.S. degree in mathematics from Soochow University, Taipei, Taiwan, in 1973, the M.S. degree in mathematics from the Institute of Mathematics, National Tsing Hua University, Hsinchu, Taiwan, in 1975, the M.A. degree in mathematics from the State University of New York, Stony Brook, NY, USA, in 1977, the M.S. and M.S.E.E. degrees from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 1982, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, MD, USA, in 1987.

He has been with the University of Maryland, Baltimore County (UMBC), Baltimore, MD, USA, since 1987, where he is a Professor with the Department of Computer Science and Electrical Engineering. He is holding the Chang Jiang Scholar Chair Professorship and is the Director of the Center for Hyperspectral Imaging in Remote Sensing (CHIRS), Dalian Maritime University, Dalian, China, since 2016. In addition, he has been a Chair Professor of National Chiao Tung University, Hsinchu, Taiwan since 2019. He holds several patents on hyperspectral image processing. He authored four books, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification* (Kluwer Academic Publishers, 2003), *Hyperspectral Data Processing: Algorithm Design and Analysis* (John Wiley & Sons, 2013), *Real Time Progressive Hyperspectral Image Processing: Endmember Finding and Anomaly Detection* (Springer, 2016), and *Recursive Hyperspectral Sample and Band Processing: Algorithm Architecture and Implementation* (Springer, 2017). In addition, he also edited two books, *Recent Advances in Hyperspectral Signal and Image Processing* (2006) and *Hyperspectral Data Exploitation: Theory and Applications* (John Wiley & Sons, 2007) and co-edited with A. Plaza a book on *High Performance Computing in Remote Sensing* (CRC Press, 2007). His research interests include multispectral/hyperspectral image processing, automatic target recognition, and medical imaging.

Dr. Chang is a fellow of Society for Photo-Optical Instrumentation Engineers (SPIE). He was a recipient of the National Research Council Senior Research Associateship Award from 2002 to 2003 sponsored by the U.S. Army Soldier and Biological Chemical Command, Edgewood Chemical and Biological Center, Aberdeen Proving Ground, Maryland.