Advanced Sciences and Technologies
for Security Applications

Editor-in-Chief: B. Javidi, University of Connecticut, Storrs, CT, USA

Editorial Board: E. Carapezza, DARPA, USA
J.P. Huignard, Thales, France
N. Nasrabadi, US Army
H. Tiziani, Stuttgart, Germany
T. Tschudi, Darmstadt, Germany
E. Watson, US Air Force, USA
T. Yatagai, Tsukuba, Japan
Advanced Sciences and Technologies for Security Applications

The series Advanced Sciences and Technologies for Security Applications focuses on research monographs in the areas of
— Recognition and identification (including optical imaging, biometrics, authentication, verification, and smart surveillance systems)
— Biological and chemical threat detection (including biosensors, aerosols, materials detection and forensics),
and
— Secure information systems (including encryption, and optical and photonic systems).
The series is intended to give an overview at the highest research level at the frontier of research in the physical sciences.
The editors encourage prospective authors to correspond with them in advance of submitting a manuscript. Submission of manuscripts should be made to the Editor-in-Chief or one of the Editors.

Editor-in-Chief
Bahram Javidi
Board of Trustees Distinguished Professor
University of Connecticut
Electrical & Computer Engineering Dept.
371 Fairfield Road, Unit 2157
Storrs, CT 06269-2157
USA
E-mail: bahram@engr.uconn.edu

Editorial Board
Edward Carapezza
Program Manager, DARPA Advanced Technology Office
3701 N. Fairfax Drive
Arlington, VA 22203-1714, USA
E-mail: ecarapezza@darpa.mil

Jean-Pierre Huignard
Thales Research & Technology
Domaine de Corbeville
91404 Orsay cedex, France
E-mail: jean-pierre.huignard@thalesgroup.com

Nasser Nasrabadi
Department of the Army
US Army Research Laboratory
ATTN: AMSRL-SE-SE
2800 Powder Mill Road
Adelphi, MD 20783-1197, USA
E-mail: nnasraba@arl.army.mil

Hans Tiziani
Universität Stuttgart
Institut für Technische Optik
Pfaffenwaldring 9
70569 Stuttgart, Germany
E-mail: tiziani@ito.uni-stuttgart.de

Theo Tschudi
Technische Universität Darmstadt
Institut für Angewandte Physik
AG Licht- und Teilchenoptik
Hochschulstraße 6
D-64289 Darmstadt
E-mail: Theo.Tschudi@physik.tu-darmstadt.de

Edward A. Watson
AFRL/SNJ Bldg 620
2241 Avionics Circle, Ste 2
WPAFB, OH 45433-7304
E-mail: Edward.Watson@wpafb.af.mil

Toyohiko Yatagai
Institute of Applied Physics
University of Tsukuba
Tsukuba Ibaraki 305, Japan
E-mail: yatagai@bukko.bk.tsukuba.ac.jp
Firooz Sadjadi:
For my mother Fakhri

Bahram Javidi:
For my wife Bethany
Contents

Preface ........................................................................................................... ix

Contributors ............................................................................................... xi

1 Kernel-Based Nonlinear Subspace Target Detection
   for Hyperspectral Imagery ................................................................. 1
   Heesung Kwon and Nasser M. Nasrabadi

2 Theory of Invariant Algebra and Its Use in
   Automatic Target Recognition ..................................................... 25
   Firooz Sadjadi

3 Automatic Recognition of Underground Targets
   Using Time–Frequency Analysis and
   Optimization Techniques ............................................................... 41
   Guillermo C. Gaunaurd and Hans C. Strifors

4 A Weighted Zak Transform, Its Properties,
   and Applications to Signal Processing ........................................... 57
   Izidor Gertner and George A. Geri

5 Using Polarization Features of Visible Light
   for Automatic Landmine Detection ................................................. 73
   Wim de Jong and John Schavemaker

6 The Physics of Polarization-Sensitive Optical Imaging .... 91
   Cornell S. L. Chun

7 Dispersion, Its Effects, and Compensation ......................... 105
   Leon Cohen and Patrick Loughlin

8 Multisensor Target Recognition in Image Response
   Space Using Evolutionary Algorithms ............................................ 127
   Igor Maslov and Izidor Gertner
Contents

9  Biophysics of the Eye in Computer Vision: Methods and Advanced Technologies ............................................. 143
    Riad I. Hammoud and Dan Witzner Hansen

10 Two Approaches to 3D Microorganism Recognition Using Single Exposure Online (SEOL) Digital Holography ...... 175
    Seokwon Yeom, Inkyu Moon, and Bahram Javidi

11 Distortion-Tolerant 3D Object Recognition by Using Single Exposure On-Axis Digital Holography .................... 195
    Daesuk Kim and Bahram Javidi

12 Design of Distortion-Invariant Optical ID Tags for Remote Identification and Verification of Objects ............ 207
    Elisabet Pérez-Cabré, María Sagrario Millán, and Bahram Javidi

13 Speckle Elimination With a Maximum Likelihood Estimation and an Isoline Regularization ........................ 227
    Nicolas Bertaux, Yann Frauel, and Bahram Javidi

Index ........................................................................................................................................... 245
Preface

This book is an attempt to provide a view of automated target recognition from the physics of the signatures. Automated target recognition (ATR) can be viewed as an inverse problem in electromagnetic and acoustics. Targets of interest are sensed by reflection/scattering, refraction or emission; the sensed signatures are then transmitted through some transmitting medium before being intercepted by detectors. The main goal of ATR is to use these signatures to classify the original objects. Our experience has indicated that only through understanding of the direct problem, the transformations that a signal goes through before being detected, one can hope of solving the inverse problem. So beside attention paid to abstract pattern classification techniques, one needs to also concentrate on the physics of the signatures, be it from a polarimetric infrared, hyperspectral, or ultra-wide band radar. Viewed in this context, ATR can be seen as an exciting physics problem whose general solution remains elusive but no too far off.

With the rapid advances in sensor designs, high speed computers and low cost platforms such as unmanned vehicles, the problem of ATR is becoming more pertinent to a much wider group of scientist and engineers than before. This book is intended to address this need by providing a cross sectional view of various issues related to the physics of the ATR.

We are grateful to the contributors, many of who are our friends and colleagues for many years for their outstanding contributions. All are imminent in their fields and most have been active in the ATR field for decades. We would like to thank the Springer Editor and staff Ms.Virginia Lipsy and Ms. Vasudha Gandi, for their support, and assistance in the preparation of this manuscript.
Contributors

N. Bertaux
Institut Fresnel - Equipe Φ-T
Unité Mixte de Recherche
6133
Domaine Universitaire de Saint Jérôme
13397 Marseille, France

C.S.L. Chun
Physics Innovations Inc.
P.O. Box 2171
Inver Grove Heights
MN 55076-81-71

L. Cohen
City University of New York
695 Park Avenue
New York, NY 10021

Dr. W. de Jong
TNO Defence, Security and Safety
Oude Waalsdorperweg 63
2597 AK The Hague
The Netherlands

Y. Frauel
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México
México, DF, Mexico

B. Gaunaud
Army Research Laboratory
2800 Powder Mill Road
Adelphi, MD 20783-1197

G.A. Geri
Link Simulation & Training
6030 S. Kent
St. Mesa, AZ 85212

I. Gertner
Computer Science Department
City College of the City University of New York
138 Street, Convent Avenue
New York, NY 10031

R.I. Hammoud
Delphi Electronics and Safety
One Corporate Center
P.O. Box 9005
Kokomo, IN 46904-9005

D.W. Hansen
IT University, Copenhagen
Rued Langgaardsvej 7
2300 Copenhagen S, Denmark

B. Javidi
Department of Electrical and Computer Engineering, U-2157
University of Connecticut
Storrs, CT 06269-2157
Contributors

D. Kim
Electrical and Computer Engineering Department
University of Connecticut
Storrs, CT 06269-2157

H. Kwon
U.S. Army Research Laboratory
ATTN: AMSRL-SE-SE
2800 Powder Mill Road
Adelphi, MD 20783

P. Loughlin
University of Pittsburgh
348 Benedum Hall
Pittsburgh, PA 15261

I. Maslov
Department of Computer Science
CUNY/Graduate Center
365 Fifth Ave.
New York, NY 10016

M.S. Millan
Department of Optics and Optometry
Universitat Politècnica de Catalunya
Violinista Vellsolà 37
08222 Terrassa, Spain

I. Moon
Department of Electrical and Computer Engineering, U-2157
University of Connecticut
Storrs, CT 06269-2157

N.M. Nasrabadi
U.S. Army Research Laboratory
ATTN: AMSRL-SE-SE
2800 Powder Mill Road
Adelphi, MD 20783

E. Perez-Cabrè
Department of Optics and Optometry
Universitat Politècnica de Catalunya
Violinista Vellsolà 37
08222 Terrassa, Spain

F. Sadjadi
Lockheed Martin Corporation
3400 High Crest Road
Saint Anthony, MN 55418

Dr. J.G.M. Schavemaker
TNO Defence, Security and Safety
Oude Waalsdorperweg 63
2597 AK The Hague
The Netherlands

H. Strifors
Swedish Defense Research Agency (FOI)
P.O. Box 1165
SE-58111 Linköping, Sweden

S. Yeom
Department of Electrical and Computer Engineering, U-2157
University of Connecticut
Storrs, CT 06269-2157
Kernel-Based Nonlinear Subspace Target Detection for Hyperspectral Imagery

Heesung Kwon and Nasser M. Nasrabadi

U.S. Army Research Laboratory ATTN: AMSRL-SE-SE 2800
Powder Mill Road Adelphi, MD 20783

Summary. In this book chapter, we compare several detection algorithms that are based on spectral matched (subspace) filters. Nonlinear (kernel) versions of these spectral matched detectors are also given and their performance is compared with the linear versions. The kernel-based matched filters and kernel matched subspace detectors exploit the nonlinear correlations between the spectral bands that is ignored by the conventional detectors. Nonlinear realization is mainly pursued to reduce data complexity in a high-dimensional feature space and consequently providing simpler decision rules for data discrimination. Several well-known matched detectors, such as matched subspace detector, orthogonal subspace detector, spectral matched filter, and adaptive subspace detector (adaptive cosine estimator), are extended to their corresponding kernel versions by using the idea of kernel-based learning theory. In kernel-based detection algorithms, the data are implicitly mapped into a high-dimensional kernel feature space by a nonlinear mapping which is associated with a kernel function. The detection algorithm is then derived in the feature space which is kernelized in terms of the kernel functions in order to avoid explicit computation in the high-dimensional feature space. Experimental results based on simulated toy-examples and real hyperspectral imagery show that the kernel versions of these detectors outperform the conventional linear detectors.

1.1 Introduction

Detecting signals of interest, particularly with wide signal variability, in noisy environments has long been a challenging issue in various fields of signal processing. Among a number of previously developed detectors, the well-known matched subspace detector (MSD) [1], orthogonal subspace detector (OSD) [1, 2], spectral matched filter (SMF) [3, 4], and adaptive subspace detectors (ASD) also known as adaptive cosine estimator (ACE) [5, 6], have been widely used to detect a desired signal (target).
Matched signal detectors, such as spectral matched filter and matched subspace detectors (whether adaptive or nonadaptive), only exploit second-order correlations, thus completely ignoring nonlinear (higher order) spectral interband correlations that could be crucial to discriminate between target and background. In this chapter, our aim is to introduce nonlinear versions of MSD, OSD, SMF and ASD detectors which effectively exploit the higher order spectral interband correlations in a high-(possibly infinite) dimensional feature space associated with a certain nonlinear mapping via kernel-based learning methods [7]. A nonlinear mapping of the input data into a high-dimensional feature space is often expected to increase the data separability and reduce the complexity of the corresponding data structure [8].

The nonlinear versions of a number of signal processing techniques such as principal component analysis (PCA) [9], Fisher discriminant analysis [10], clustering in feature space [8], linear classifiers [11], nonlinear feature extraction based on kernel orthogonal centroid method [12], kernel-matched signal detectors for target detection [13, 14, 15], kernel-based anomaly detection [16], classification in the kernel-based nonlinear subspace [17], and classifiers based on kernel Bayes rule [18] have already been defined in kernel space. In [19] kernels were used as generalized dissimilarity measures for classification and in [20] kernel methods were applied to face recognition.

This chapter is organized as follows. Section 1.2 provides the background to the kernel-based learning methods and kernel trick. Section 1.3 introduces linear matched subspace and its kernel version. Orthogonal subspace detector is defined in section 1.4 as well as its kernel version. In Section 1.5, we describe the conventional spectral matched filter and its kernel version in the feature space and reformulate the the expression in terms of the kernel function using the kernel trick. Finally, in Section 1.6 the adaptive subspace detector and its kernel version are introduced. Performance comparison between the conventional and the kernel version of these algorithms is provided in Section 1.7. Conclusions and comparisons of these algorithms are given in Section 1.8.

### 1.2 Kernel Methods and Kernel Trick

Suppose that the input hyperspectral data are represented by the data space \( \mathcal{X} \subseteq \mathbb{R}^l \) and \( \mathcal{F} \) is a feature space associated with \( \mathcal{X} \) by a nonlinear mapping function \( \Phi \)

\[
\Phi : \mathcal{X} \rightarrow \mathcal{F}, \mathbf{x} \mapsto \Phi(\mathbf{x}),
\]

(1.1)

where \( \mathbf{x} \) is an input vector in \( \mathcal{X} \) which is mapped into a potentially much higher—(could be infinite)—dimensional feature space. Because of the high dimensionality of the feature space \( \mathcal{F} \), it is computationally not feasible to implement any algorithm directly in the feature space. However, kernel-based learning algorithms use an effective kernel trick given by Eq. (1.2) to implement dot products in the feature space by employing kernel functions [7].
idea in kernel-based techniques is to obtain a nonlinear version of an algorithm defined in the input space by implicitly redefining it in the feature space and then converting it in terms of dot products. The kernel trick is then used to implicitly compute the dot products in \( \mathcal{F} \) without mapping the input vectors into \( \mathcal{F} \); therefore, in the kernel methods, the mapping \( \Phi \) does not need to be identified.

The kernel representation for the dot products in \( \mathcal{F} \) is expressed as

\[
k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j),
\]

where \( k \) is a kernel function in terms of the original data. There are a large number of Mercer kernels that have the kernel trick property; see [7] for detailed information about the properties of different kernels and kernel-based learning. Our choice of kernel in this chapter is the Gaussian RBF kernel and the associated nonlinear function \( \Phi \) with this kernel generates a feature space of infinite dimensionality.

### 1.3 Linear Matched Subspace Detector and Kernel Matched Subspace Detector

In this model the target pixel vectors are expressed as a linear combination of target spectral signature and background spectral signature, which are represented by subspace target spectra and subspace background spectra, respectively. The hyperspectral target detection problem in a \( p \)-dimensional input space is expressed as two competing hypotheses \( H_0 \) and \( H_1 \)

\[
H_0 : y = B\zeta + n, \quad \text{Target absent (1.3)}
\]

\[
H_1 : y = T\theta + B\zeta + n = [TB]\begin{bmatrix} \theta \\ \zeta \end{bmatrix} + n, \quad \text{Target present}
\]

where \( T \) and \( B \) represent orthogonal matrices whose \( p \)-dimensional column vectors span the target and background subspaces, respectively; \( \theta \) and \( \zeta \) are unknown vectors whose entries are coefficients that account for the abundances of the corresponding column vectors of \( T \) and \( B \), respectively; \( n \) represents Gaussian random noise \( (n \in \mathcal{N}(0, \sigma^2 I)) \); and \( [TB] \) is a concatenated matrix of \( T \) and \( B \). The numbers of the column vectors of \( T \) and \( B \), \( N_t \) and \( N_b \), respectively, are usually smaller than \( p \) \((N_t, N_b < p)\).

#### 1.3.1 Generalized Likelihood Ratio Test (GLRT) for Target Detection

Given the linear subspace detection model and the two hypotheses about how the input vector is generated, as shown by (1.3), the likelihood ratio test (LRT) is used to predict whether the input vector \( y \) includes the target and
is defined by
\[ l(y) = \frac{p_1(y|H_1)}{p_0(y|H_0)} \overset{H_1}{\gtrless} \eta, \tag{1.4} \]
where \( p_0(y|H_0) \) and \( p_1(y|H_1) \) represent the class conditional probability densities of \( y \) given the hypotheses \( H_0 \) and \( H_1 \), respectively, and \( \eta \) is a threshold of the test. \( p_0(y|H_0) \) and \( p_1(y|H_1) \) can be expressed as Gaussian probability densities \( N(B\zeta,\sigma^2I) \) and \( N(T\theta + B\zeta,\sigma^2I) \), respectively, since \( n \) are assumed to be Gaussian random noise. The LRT is derived from the Neyman–Pearson criterion where the probability detection \( P_D \) is maximized while the probability of false alarm \( P_F \) is kept a constant [21]. \( l(y) \) is compared to \( \eta \) to make a final decision about which hypothesis best relates to \( y \).

The LRT includes the unknown parameters \( \zeta \) and \( \theta \) that need to be estimated using the maximum likelihood principle. The generalized likelihood ratio test (GLRT) is directly obtained from \( l(y) \) by replacing the unknown parameters with their maximum likelihood estimates (MLEs) \( \hat{\zeta} \) and \( \hat{\theta} \) and by taking \((P/2)\)-root
\[ L_2(y) = (\hat{l}(y))^{2/P} = \frac{y^T(I - PB)y}{y^T(I - P_{TB})y} \overset{H_1}{\gtrless} \eta, \tag{1.5} \]
where \( PB = B(B^TB)^{-1}B^T = BB^T \) is a projection matrix associated with the \( N_b \)-dimensional background subspace \(<B>\); \( P_{TB} \) is a projection matrix associated with the \((N_b = N_b + N_t)\)-dimensional target-and-background subspace \(<TB>\)
\[ P_{TB} = [T B] [D T B]^{-1} [T B]^T. \tag{1.6} \]

### 1.3.2 Linear Subspace Models Defined in the Feature Space \( \mathcal{F} \)

The hyperspectral detection problem based on the target and background subspaces can be described in the feature space \( \mathcal{F} \) as
\[ H_{0\phi} : \Phi(y) = B\phi\zeta\phi + n_\phi, \quad \text{Target absent} \]
\[ H_{1\phi} : \Phi(y) = T\phi\theta\phi + B\phi\zeta\phi + n_\phi = [T\phi B\phi] \begin{bmatrix} \theta\phi \\ \zeta\phi \end{bmatrix} + n_\phi, \quad \text{Target present} \tag{1.7} \]
where \( T\phi \) and \( B\phi \) represent full-rank matrices whose column vectors span target and background subspaces \(<B\phi>\) and \(<T\phi>\) in \( \mathcal{F} \), respectively; \( \theta\phi \) and \( \zeta\phi \) are unknown vectors whose entries are coefficients that account for the abundances of the corresponding column vectors of \( T\phi \) and \( B\phi \), respectively; \( n_\phi \) represents Gaussian random noise; and \([T\phi B\phi]\) is a concatenated matrix of \( T\phi \) and \( B\phi \). In general, any sets of basis vectors that span the corresponding subspace can be used as the column vectors of \( T\phi \) and \( B\phi \). In the proposed method we use the significant eigenvectors of the target and
background covariance matrices (\(C_{T_\phi}\) and \(C_{B_\phi}\)) in \(F\) as the the column vectors of \(T_\phi\) and \(B_\phi\), respectively. \(C_{T_\phi}\) and \(C_{B_\phi}\) are based on the zero mean (centered) target and background sample sets (\(Z_T\) and \(Z_B\)), respectively:

\[
C_{T_\phi} = \frac{1}{N} \sum_{i=1}^{N} \phi(y_i)\phi(y_i)^T, \quad \text{for } y_i \in Z_T,
\]

\[
C_{B_\phi} = \frac{1}{M} \sum_{i=1}^{M} \phi(y_i)\phi(y_i)^T, \quad \text{for } y_i \in Z_B,
\]

where \(N\) and \(M\) represent the number of centered training samples in \(Z_T\) and \(Z_B\), respectively.

Using a similar reasoning as described in the previous subsection, the GLRT of the hyperspectral detection problem depicted by the model in (1.7) is given by

\[
L_2(\phi(y)) = \frac{\phi(y)^T(P_{I_\phi} - P_{B_\phi})\phi(y)}{\phi(y)^T(P_{I_\phi} - P_{T_\phi B_\phi})\phi(y)},
\]

(1.8)

where \(P_{I_\phi}\) represents an identity projection operator in \(F\); \(P_{B_\phi} = B_\phi (B_\phi^T B_\phi)^{-1} B_\phi^T\) is a background projection matrix; and \(P_{T_\phi B_\phi}\) is a joint target-and-background projection matrix in \(F\)

\[
P_{T_\phi B_\phi} = [T_\phi \ B_\phi] \left[ [T_\phi \ B_\phi]^T \ [T_\phi \ B_\phi] \right]^{-1} \left[ T_\phi \ B_\phi \right]^T
\]

(1.9)

\[
= [T_\phi \ B_\phi] \left[ T_\phi^T T_\phi \ T_\phi^T B_\phi \ B_\phi^T \ B_\phi^2 \ B_\phi \right]^{-1} \left[ T_\phi^T \ B_\phi \right].
\]

1.3.3 Kernelizing MSD in Feature Space

To kernelize (1.8) we will separately kernelize the numerator and the denominator. First consider its numerator,

\[
\phi(y)^T(P_{I_\phi} - P_{B_\phi})\phi(y)\phi(y)^T P_{I_\phi} \phi(y) - \phi(y)^T B_\phi B_\phi^2 \phi(y).
\]

Using (1.59), as shown in Appendix I, \(B_\phi\) and \(T_\phi\) can be written in terms of their corresponding data spaces as

\[
B_\phi = \begin{bmatrix} e_b^1 & e_b^2 & \ldots & e_b^{N_b} \end{bmatrix} = \phi_{Z_B} \tilde{B},
\]

(1.11)

\[
T_\phi = \begin{bmatrix} e_t^1 & e_t^2 & \ldots & e_t^{N_t} \end{bmatrix} = \phi_{Z_T} \tilde{T},
\]

(1.12)

where \(e_b^i\) and \(e_t^i\) are the significant eigenvectors of \(C_{B_\phi}\) and \(C_{T_\phi}\), respectively; \(\phi_{Z_B} = [\phi(y_1) \phi(y_2) \ldots \phi(y_M)], y_i \in Z_B\) and \(\phi_{Z_T} = [\phi(y_1) \phi(y_2) \ldots \phi(y_N)], y_i \in Z_T\); the column vectors of \(\tilde{B}\) and \(\tilde{T}\) represent only the significant normalized eigenvectors (\(\tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_{N_b}\)) and (\(\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_{N_t}\)) of the background centered kernel matrix \(K(Z_B, Z_B) = (K)_{ij} = k(y_i, y_j), y_i, y_j \in Z_B\) and target.
centered kernel matrix $K(Z_T, Z_T) = (K)_{ij} = k(y_i, y_j)$, $y_i, y_j \in Z_T$, respectively.

Using (1.11) the projection of $\Phi(y)$ onto $B_\phi$ becomes

$$B_\phi^T \Phi(y) = \left[ e_b^1 \ e_b^2 \ \cdots \ e_b^{N_b} \right]^T \Phi(y) = \begin{bmatrix} \tilde{\beta}_1^T \Phi_{Z_b} \Phi(y) \\ \tilde{\beta}_2^T \Phi_{Z_b} \Phi(y) \\ \vdots \\ \tilde{\beta}_{N_b}^T \Phi_{Z_b} \Phi(y) \end{bmatrix} = \tilde{B}^T k(Z_B, y), \quad (1.13)$$

and, similarly, using (1.12) the projection onto $T_\phi$ is

$$T_\phi^T \Phi(y) = \left[ e_t^1 \ e_t^2 \ \cdots \ e_t^{N_t} \right]^T \Phi(y) = \begin{bmatrix} \tilde{\alpha}_1^T \Phi_{Z_t} \Phi(y) \\ \tilde{\alpha}_2^T \Phi_{Z_t} \Phi(y) \\ \vdots \\ \tilde{\alpha}_{N_t}^T \Phi_{Z_t} \Phi(y) \end{bmatrix} = \tilde{T}^T k(Z_T, y), \quad (1.14)$$

where $k(Z_B, y)$ and $k(Z_T, y)$, referred to as the empirical kernel maps in the machine learning literature [7], are column vectors whose entries are $k(x_i, y)$ for $x_i \in Z_B$ and $x_i \in Z_T$, respectively. Now using (1.13) we can write

$$\Phi(y)^T \tilde{B}_\phi \tilde{B}_\phi^T \Phi(y) = k(Z_B, y)^T \tilde{B} \tilde{B}^T k(Z_B, y). \quad (1.15)$$

The projection onto the identity operator $\Phi(y)^T P_{L_\phi} \Phi(y)$ also needs to be kernelized. $P_{L_\phi}$ is defined as $P_{L_\phi} := \Omega_\phi \Omega_\phi^T$, where $\Omega_\phi = \left[ e_q^1 \ e_q^2 \ \cdots \right]$ is a matrix whose columns are all the eigenvectors with $\lambda \neq 0$ that are in the span of $\Phi(y_i)$, $y_i \in Z_T \cup Z_B := Z_{TB}$. From (1.59) $\Omega_\phi$ can be expressed as

$$\Omega_\phi = \left[ e_q^1 \ e_q^2 \ \cdots \ e_q^{N_{eq}} \right] = \Phi_{Z_{tn}} \tilde{\Delta}, \quad (1.16)$$

where $\Phi_{Z_{tn}} = \Phi_{Z_T} \cup \Phi_{Z_B}$ and $\tilde{\Delta}$ is a matrix whose columns are the eigenvectors ($\kappa_1, \kappa_2, \cdots, \kappa_{N_{eq}}$) of the centered kernel matrix $K(Z_{TB}, Z_{TB}) = (K)_{ij} = k(y_i, y_j)$, $y_i, y_j \in Z_{TB}$ with nonzero eigenvalues, normalized by the square root of their associated eigenvalues. Using $P_{L_\phi} = \Omega_\phi \Omega_\phi^T$ and (1.16)

$$\Phi(y)^T P_{L_\phi} \Phi(y) = \Phi(y)^T \Phi_{Z_{tn}} \tilde{\Delta} \tilde{\Delta}^T \Phi_{Z_{tn}} \Phi(y) = k(Z_{TB}, y)^T \tilde{\Delta} \tilde{\Delta}^T k(Z_{TB}, y), \quad (1.17)$$

where $k(Z_{TB}, y)$ is a concatenated vector $[k(Z_T, y)^T \ k(Z_B, y)^T]^T$. The kernelized numerator of (1.8) is now given by

$$k(Z_{TB}, y)^T \tilde{\Delta} \tilde{\Delta}^T k(Z_{TB}, y) - k(Z_B, y)^T \tilde{B} \tilde{B}^T K(Z_B, y). \quad (1.18)$$
We now kernelize $\Phi(y)^T P_{T_y B_y} \Phi(y)$ in the denominator of (1.8) to complete the kernelization process. Using (1.9), (1.11) and (1.12)

$$
\Phi(y)^T P_{T_y B_y} \Phi(y)
$$

$$
= \Phi(y)^T \left[ T_\Phi B_\Phi \right]^{-1} \left[ T_\Phi^T \right] \Phi(y)
$$

$$
= \left[ K(Z_T, y)^T T \right]^{-1} \left[ T^T K(Z_T, Z_T) T \right]^{-1}
$$

$$
\times \left[ B^T K(Z_B, y) \right].
$$

Finally, substituting (1.15), (1.17) and (1.19) into (1.8) the kernelized GLRT is given by

$$
L_{2K} = \left( k(Z_{TB}, y)^T \tilde{T} K(Z_{TB}, y) - k(Z_B, y)^T \tilde{B} \tilde{T} \tilde{B}^T k(Z_B, y) \right) / \left( k(Z_{TB}, y)^T \tilde{T} K(Z_{TB}, y) - \left[ k(Z_T, y)^T \tilde{T} k(Z_B, y)^T \tilde{B} \right] \Lambda_1^{-1}
$$

$$
\times \left[ \tilde{B}^T k(Z_B, y) \right],
$$

(1.20)

where

$$
\Lambda_1 = \left[ \tilde{T}^T K(Z_T, Z_T) \tilde{T} \tilde{B}^T K(Z_B, Z_B) \tilde{B} \right].
$$

In the above derivation (1.20) we assumed that the mapped input data were centered in the feature space $\Phi_c(x_i) = \Phi(x_i) - \mu_{b_\phi}$, where $\mu_{b_\phi}$ represents the estimated mean in the feature space given by $\mu_{b_\phi} = \frac{1}{N} \sum_{i=1}^{N} \Phi(x_i)$. However, the original data are usually not centered and the estimated mean in the feature space cannot be explicitly computed, therefore, the kernel matrices have to be properly centered as shown in (1.70). The empirical kernel maps $k(Z_T, y)$, $k(Z_B, y)$, and $k(Z_{TB}, y)$ have to be centered by removing their corresponding empirical kernel mean (e.g., $\tilde{k}(Z_T, y) = k(Z_T, y) - \frac{1}{N} \sum_{i=1}^{N} k(y_i, y), y_i \in Z_T$).

1.4 OSP and Kernel OSP Algorithms

1.4.1 Linear Spectral Mixture Model

The OSP algorithm [2] is based on maximizing the SNR (signal-to-noise ratio) in the subspace orthogonal to the background subspace and depends only on
the noise second-order statistics. It also does not provide an estimate of the abundance measure for the desired end member in the mixed pixel. In [22] it was shown that the OSP classifier is related to the unconstrained least-squares estimate or the MLE (similarly derived by [1]) of the unknown signature abundance by a scaling factor.

A linear mixture model for pixel \( y \) consisting of \( p \) spectral bands is described by
\[
y = M\alpha + n, \tag{1.21}
\]
where the \((p \times l)\) matrix \( M \) represents \( l \) endmembers spectra, \( \alpha \) is a \((l \times p)\) column vector whose elements are the coefficients that account for the proportions (abundances) of each endmember spectrum contributing to the mixed pixel, and \( n \) is an \((p \times p)\) vector representing an additive zero-mean Gaussian noise with covariance matrix \( \sigma I \) and \( I \) is the \((l \times l)\) identity matrix. Assuming now we want to identify one particular signature (e.g., a military target) with a given spectral signature \( d \) and a corresponding abundance measure \( \alpha_l \), we can represent \( M \) and \( \alpha \) in partition form as \( M = (U : d) \) and \( \alpha = [\gamma \alpha_l] \) then model (1.21) can be rewritten as
\[
r = d\alpha_l + B\gamma + n, \tag{1.22}
\]
where the columns of \( B \) represent the undesired spectral signatures (background signatures or eigenvectors) and the column vector \( \gamma \) is the abundance measures for the undesired spectral signatures. The reason for rewriting model (1.21) as (1.22) is to separate \( B \) from \( M \) in order to show how to annihilate \( B \) from an observed input pixel prior to classification.

To remove the undesired signature, the background rejection operator is given by the \((l \times l)\) matrix
\[
P_B^\perp = I - BB^\# \tag{1.23}
\]
where \( B^\# = (B^TB)^{-1}B^T \) is the pseudoinverse of \( B \). Applying \( P_B^\perp \) to the model (1.22) results in
\[
P_B^\perp r = P_B^\perp d\alpha_l + P_B^\perp n. \tag{1.24}
\]
The operator \( w \) that maximizes the SNR of the filter output \( wP_B^\perp y \),
\[
\text{SNR}(w) = \frac{w^TP_B^\perp d|\alpha_l|^2|d^TP_B^\perp w|}{w^TP_B^\perp E|nn^T|P_B^\perp w}, \tag{1.25}
\]
as shown in [2], is given by the matched filter \( w = \kappa d \), where \( \kappa \) is a constant. The OSP operator is now given by
\[
q_{OSP}^T = d^TP_B^\perp \tag{1.26}
\]
which consists of a background signature rejecter followed by a matched filter. The output of the OSP classifier is given by

$$D_{OSP} = q_{OSP}^{T} r = d^{T} P_{B}^{+} y.$$  

(1.27)

### 1.4.2 OSP in Feature Space and Its Kernel Version

The mixture model in the high dimensional feature space $F$ is given by

$$\Phi(r) = M_{\Phi} \alpha_{\Phi} + n_{\Phi},$$  

(1.28)

where $M_{\Phi}$ is a matrix whose columns are the endmembers spectra in the feature space; $\alpha_{\Phi}$ is a coefficient vector that accounts for the abundances of each endmember spectrum in the feature space; $n_{\Phi}$ is an additive zero-mean noise. The model (1.28) can also be rewritten as

$$\Phi(r) = \Phi(d) \alpha_{p_{\Phi}} + B_{\Phi} \gamma_{\Phi} + n_{\Phi},$$  

(1.29)

where $\Phi(d)$ represent the spectral signature of the desired target in the feature space with the corresponding abundance $\alpha_{p_{\Phi}}$ and the columns of $B_{\Phi}$ represent the undesired background signatures in the feature space which are implemented by the eigenvectors of the background covariance matrix.

The output of the OSP classifier in the feature space is given by

$$D_{OSP_{\Phi}} = q_{OSP_{\Phi}}^{T} r = \Phi(d)^{T} (I_{\Phi} - B_{\Phi} B_{\Phi}^{T}) \Phi(r)$$  

(1.30)

where $I_{\Phi}$ is the identity matrix in the feature space. This output (1.30) is very similar to the numerator of (1.8). It can easily be shown that the kernelized version of (1.30) is now given by

$$D_{KOSP} = k(Z_{Bd},d)^{T} \tilde{Y}^{T} k(Z_{Bd},y) - k(Z_{B},y)^{T} B \tilde{B} B^{T} k(Z_{B},y)$$  

(1.31)

where $Z_{B} = [x_{1}, x_{2}, \ldots, x_{N}]$ correspond to $N$ input background spectral signatures and $\tilde{B} = (\tilde{\beta}_{1}, \tilde{\beta}_{2}, \ldots, \tilde{\beta}_{N_{b}})^{T}$ are the $N_{b}$ significant eigenvectors of the centered kernel matrix (Gram matrix) $K(Z_{B},Z_{B})$ normalized by the square root of their corresponding eigenvalues. $k(Z_{B},r)$ and $k(Z_{B},d)$ are column vectors whose entries are $k(x_{i},y)$ and $k(x_{i},d)$ for $x_{i} \in Z_{B}$, respectively. $Z_{Bd} = Z_{B} \cup d$ and $\tilde{Y}$ is a matrix whose columns are the $N_{bd}$ eigenvectors $(v_{1}, v_{2}, \ldots, v_{N_{bd}})$ of the centered kernel matrix $K(Z_{Bd},Z_{Bd}) = (K)_{ij} = k(x_{i},x_{j})$, $x_{i}, x_{j} \in Z_{B} \cup d$ with nonzero eigenvalues, normalized by the square root of their associated eigenvalues. Also $k(Z_{Bd},y)$ is the concatenated vector $[k(Z_{B},r)^{T} k(d,y)^{T}]^{T}$ and $k(Z_{Bd},d)$ is the concatenated vector $[k(Z_{Bd},d)^{T} k(d,d)^{T}]^{T}$. In the above derivation (1.31) we assumed that the mapped input data were centered in the feature space. For non-centered data the kernel matrices and the empirical kernel maps have to be properly centered as was shown in the previous Section 1.3.3.
1.5 Linear Spectral Matched Filter and Kernel Spectral Matched Filter

1.5.1 Linear Spectral Matched Filter

In this section, we introduce the concept of linear SMF. The constrained least squares approach is used to derive the linear SMF. Let the input spectral signal \( x \) be \( x = [x(1), x(2), \ldots, x(p)]^T \) consisting of \( p \) spectral bands. We can model each spectral observation as a linear combination of the target spectral signature and noise

\[
x = as + n, \tag{1.32}
\]

where \( a \) is an attenuation constant (target abundance measure). When \( a = 0 \) no target is present and when \( a > 0 \) target is present, vector \( s = [s(1), s(2), \ldots, s(p)]^T \) contains the spectral signature of the target and vector \( n \) contains the added background clutter noise.

Let us define \( X \) to be a \( p \times N \) matrix of the \( N \) mean-removed background reference pixels (centered) obtained from the input image. Let each observation spectral pixel to be represented as a column in the sample matrix \( X \)

\[
X = [x_1, x_2, \ldots, x_N]. \tag{1.33}
\]

We can design a linear matched filter \( w = [w(1), w(2), \ldots, w(p)]^T \) such that the desired target signal \( s \) is passed through while the average filter output energy is minimized. The solution to this constrained least squares minimization problem is given by

\[
w = \frac{\hat{R}^{-1}s}{s^T \hat{R}^{-1}s} \tag{1.34}
\]

where \( \hat{R} \) represents the estimated correlation matrix for the reference data. The above expression is referred to as Minimum Variance Distortionless Response (MVDR) beamformer in the array processing literature [23, 24] and more recently the same expression was also obtained in [25] for hyperspectral target detection and was called Constrained Energy Minimization (CEM) filter. The output of the linear filter for the test input \( r \), given the estimated correlation matrix is given by

\[
y_r = w^Tr = \frac{s^T \hat{R}^{-1}r}{s^T \hat{R}^{-1}s}. \tag{1.35}
\]

If the observation data is centered a similar expression is obtained for the centered data which is given by

\[
y_r = w^Tr = \frac{s^T \hat{C}^{-1}r}{s^T \hat{C}^{-1}s} \tag{1.36}
\]
where \( \hat{C} \) represents the estimated covariance matrix for the reference centered data. Similarly, in [4] and [5] it was shown that using the GLRT a similar expression as in MVDR or CEM, (1.36), can be obtained if the \( n \) is assumed to be the background Gaussian random noise distributed as \( \mathcal{N}(0, C) \) where \( C \) is the expected covariance matrix of only the background noise. This filter is referred to as matched filter in the signal processing literature or Capon method [26] in the array processing literature. In this book chapter, we implemented the matched filter given by the expression (1.36).

### 1.5.2 Spectral Matched Filter in Feature Space and Its Kernel Version

We now consider a linear model in the kernel feature space which has an equivalent nonlinear model in the original input space

\[
\phi(x) = a_\phi \phi(s) + n_\phi, \tag{1.37}
\]

where \( \phi \) is the non-linear mapping that maps the input data into a kernel feature space, \( a_\phi \) is an attenuation constant (abundance measure), the high-dimensional vector \( \phi(s) \) contains the spectral signature of the target in the feature space, and vector \( n_\phi \) contains the added noise in the feature space.

Using the constrained least squares approach that was explained in the previous section it can easily be shown that the equivalent matched filter \( w_\phi \) in the feature space is given by

\[
w_\phi = \frac{\hat{R}_\phi^{-1}\phi(s)}{\phi(s)^T \hat{R}_\phi^{-1}\phi(s)}, \tag{1.38}
\]

where \( \hat{R}_\phi \) is the estimated correlation matrix in the feature space. The correlation matrix is given by

\[
\hat{R}_\phi = \frac{1}{N} X_\phi X_\phi^T \tag{1.39}
\]

where \( X_\phi = [\phi(x_1) \phi(x_2) \cdots \phi(x_N)] \) is a matrix whose columns are the mapped input reference data in the feature space. The matched filter in the feature space (1.38) is equivalent to a non-linear matched filter in the input space and its output for the input \( \phi(r) \) is given by

\[
y_{\phi(r)} = w_\phi^T \phi(r) = \frac{\phi(s)^T \hat{R}_\phi^{-1} \phi(r)}{\phi(s)^T \hat{R}_\phi^{-1} \phi(s)}. \tag{1.40}
\]

If the data was centered the matched filter for the centered data in the feature space would be

\[
y_{\phi(r)} = w_\phi^T \phi(r) = \frac{\phi(s)^T \hat{C}_\phi^{-1} \phi(r)}{\phi(s)^T \hat{C}_\phi^{-1} \phi(s)}. \tag{1.41}
\]
We now show how to kernelize the matched filter expression (1.41) where the resulting non-linear matched filter is called the kernel matched filter. It was shown in Appendix I the pseudoinverse (inverse) of the estimated background covariance matrix can be written as

\[
\hat{C}_\phi^# = X_\phi B \Lambda^{-2} B^T X_\phi^T
\]  

(1.42)

Inserting Equation (1.42) into (1.41), it can be rewritten as

\[
y_{\phi(r)} = \frac{\Phi(s)^T X_\phi B \Lambda^{-1} B^T X_\phi^T \Phi(r)}{\Phi(s)^T X_\phi B \Lambda^{-1} B^T X_\phi^T \Phi(s)}.
\]  

(1.43)

Also using the properties of the Kernel PCA as shown in Appendix I, we have the relationship

\[
K^{-2} = \frac{1}{N^2} B \Lambda^{-2} B^T.
\]  

(1.44)

We denote \( K = K(X, X) = (K)_{ij} \) an \( N \times N \) Gram kernel matrix whose entries are the dot products \( <\Phi(x_i), \Phi(x_j)> \). Substituting (1.44) into (1.43), the kernelized version of SMF is given by

\[
y_{kr} = \frac{k(X, s)^T K^{-2} k(X, r)}{k(X, s)^T K^{-2} k(X, s)} = \frac{k_r^T K^{-2} k_r}{k_s^T K^{-2} k_s}
\]  

(1.45)

where \( k_s = k(X, s) \) and \( k_r = k(X, r) \) are the empirical kernel maps for \( s \) and \( r \), respectively. As in the previous section, the kernel matrix \( K \) as well as the empirical kernel maps, \( k_s \) and \( k_r \) needs to be properly centered.

### 1.6 Adaptive Subspace Detector and Kernel Adaptive Subspace Detector

#### 1.6.1 Linear Adaptive Subspace Detector

In this section, the GLRT under the two competing hypotheses \((H_0 \quad \text{and} \quad H_1)\) for a certain mixture model is described. The subpixel detection model for a measurement \( x \) (a pixel vector) is expressed as

\[
H_0 : x = n, \quad \text{Target absent} \\
H_1 : x = U \theta + \sigma n, \quad \text{Target present}
\]  

(1.46)

where \( U \) represents an orthogonal matrix whose column vectors are the eigenvectors that span the target subspace \( <U> \); \( \theta \) is an unknown vector whose entries are coefficients that account for the abundances of the corresponding column vectors of \( U \); \( n \) represents Gaussian random noise distributed as \( \mathcal{N}(0, C) \).

In model, \( x \) is assumed to be a background noise under \( H_0 \) and a linear combination of a target subspace signal and a scaled background noise,
distributed as \( N(U\theta, \sigma^2C) \), under \( H_1 \). The background noise under the two hypotheses is represented by the same covariance but different variances because of the existence of subpixel targets under \( H_1 \). The GLRT for the subpixel problem described in [5], (so-called ASD), is given by

\[
D_{\text{ASD}}(x) = \frac{x^T\hat{C}^{-1}U(U^T\hat{C}^{-1}U)^{-1}U^T\hat{C}^{-1}x}{x^T\hat{C}^{-1}x} \gtrless \eta_{\text{ASD}},
\]

where \( \hat{C} \) is the MLE of the covariance \( C \) and \( \eta_{\text{ASD}} \) represents a threshold. Expression (1.47) has a constant false alarm rate (CFAR) property and is also referred to as the adaptive cosine estimator because (1.47) measures the angle between \( \tilde{x} \) and \( \langle \tilde{U} \rangle \), where \( \tilde{x} = \hat{C}^{-1/2}x \) and \( \tilde{U} = \hat{C}^{-1/2}U \).

### 1.6.2 ASD in the Feature Space and Its Kernel Version

We define a new subpixel model by assuming that the input data have been implicitly mapped by a nonlinear function \( \Phi \) into a high-dimensional feature space \( F \). The subpixel model in \( F \) is then given by

\[
\begin{align*}
H_{0\Phi} : \Phi(x) &= n_{\Phi}, & \text{Target absent} \\
H_{1\Phi} : \Phi(x) &= U\Phi\theta_{\Phi} + \sigma_{\Phi}n_{\Phi}, & \text{Target present}
\end{align*}
\]

where \( U_{\Phi} \) represents a full-rank matrix whose \( M_1 \) column vectors are the eigenvectors that span target subspace \( \langle U_{\Phi} \rangle \) in \( F \); \( \theta_{\Phi} \) is unknown vectors whose entries are coefficients that account for the abundances of the corresponding column vectors of \( U_{\Phi} \); \( n_{\Phi} \) represents Gaussian random noise distributed by \( N(0, C_{\Phi}) \); and \( \sigma_{\Phi} \) is the noise variance under \( H_{1\Phi} \). The GLRT for the model (1.48) in \( F \) is now given by

\[
D(\Phi(x)) = \frac{\Phi(x)^T\hat{C}_{\Phi}^{-1}U_{\Phi}(U_{\Phi}^T\hat{C}_{\Phi}^{-1}U_{\Phi})^{-1}U_{\Phi}^T\hat{C}_{\Phi}^{-1}\Phi(x)}{\Phi(x)^T\hat{C}_{\Phi}^{-1}\Phi(x)},
\]

where \( \hat{C}_{\Phi} \) is the MLE of \( C_{\Phi} \).

We now show how to kernelize the GLRT expression (1.49) in the feature space. The inverse (pseudoinverse) background covariance matrix in (1.49) can be represented by its eigenvector decomposition (see Appendix I) given by the expression

\[
\hat{C}_{\Phi}^\# = X_{\Phi}B\Lambda^{-2}B^TX_{\Phi}^T,
\]

where \( X_{\Phi} = [\Phi_c(x_1) \Phi_c(x_2) \cdots \Phi_c(x_N)] \) represents the centered vectors in the feature space corresponding to \( N \) independent background spectral signatures \( X = [x_1 \ x_2 \ \cdots \ x_N] \) and \( B = [\beta^1 \ \beta^2 \ \cdots \ \beta^N] \) are the nonzero eigenvectors of the centered kernel matrix (Gram matrix) \( K(X, X) \) normalized by the square root of their corresponding eigenvalues. Similarly, \( U_{\Phi} \) is given by

\[
U_{\Phi} = Y_{\Phi}\hat{T},
\]
where \( \Phi = [\Phi_c(y_1) \Phi_c(y_2) \ldots \Phi_c(y_M)] \) are the centered vectors in the feature space corresponding to the \( M \) independent target spectral signatures \( Y = [y_1 y_2 \ldots y_M] \) and \( \tilde{T} = [\tilde{\alpha}^1 \tilde{\alpha}^2 \ldots \tilde{\alpha}^{M_1}] \), \( M_1 < M \), is a matrix consisting of the \( M_1 \) eigenvectors of the kernel matrix \( K(Y, Y) \) normalized by the square root of their corresponding eigenvalues. Now, the term \( \Phi(x)^T \tilde{C}_\Phi^{-1} U \Phi \) in the numerator of (1.49) becomes

\[
\Phi(x)^T \tilde{C}_\Phi^{-1} U \Phi = \Phi(x)^T X \Lambda^{-2} B^T X \Phi Y \tilde{T}^T
\]

(1.52)

where \( B \Lambda^{-2} B^T \) is replaced by \( K(X, X)^{-2} \) using (1.69), as shown in Appendix I. Similarly,

\[
U \Phi^T \tilde{C}_\Phi^{-1} U \Phi = \tilde{T}^T K(X, Y)^T K(X, X)^{-2} k(x, X) = K_x^T
\]

(1.53)

and

\[
U \Phi^T \tilde{C}_\Phi^{-1} U \Phi = \tilde{T}^T K(X, Y)^T K(X, X)^{-2} K(X, Y) \tilde{T}
\]

(1.54)

The denominator of (1.49) is also expressed as

\[
\Phi(x)^T \tilde{C}_\Phi^{-1} \Phi(x) = k(x, X)^T K(X, X)^{-2} k(x, X).
\]

(1.55)

Finally, the kernelized expression of (1.49) is given by

\[
D_{\text{KASD}}(x) = \frac{K_x^T \tilde{T} K(X, Y)^T K(X, X)^{-2} K(X, Y) \tilde{T}^T}{k(x, X)^T K(X, X)^{-2} k(x, X)}
\]

(1.56)

As in the previous sections all the kernel matrices \( K(X, Y) \) and \( K(X, X) \) need to be properly centered.

### 1.7 Experimental Results

In this section, the kernel-based matched signal detectors, such as the kernel MSD (KMSD), kernel ASD (KASD), kernel OSP (KOSP), and kernel SMF (KSMF), as well as the corresponding conventional detectors are implemented on the basis of two different types of data sets—illustrative toy data sets and real hyperspectral images that contain military targets. The Gaussian RBF kernel, \( k(x, y) = \exp(-\|x-y\|^2) \), was used to implement the kernel-based detectors. \( c \) represents the width of the Gaussian distribution, and the value of \( c \) was chosen such that the overall data variations can be fully exploited by the Gaussian RBF function; the values of \( c \) were determined experimentally.

#### 1.7.1 Illustrative Toy Examples

Figures 1.1 and 1.2 show contour and surface plots of the conventional detectors and the kernel-based detectors, on two different types of two-dimensional
Figure 1.1. Contour and surface plots of the conventional matched signal detectors and their corresponding kernel versions on a toy data set (a mixture of Gaussian).
Figure 1.2. Contour and surface plots of the conventional matched signal detectors and their corresponding kernel versions on a toy data set: in this toy example, the Gaussian mixture data shown in Fig. 1 was modified to generate nonlinearly mixed data.
toy data sets: a Gaussian mixture in Fig. 1.1 and nonlinearly mapped data in Fig. 1.2. In the contour and surface plots, data points for the desired target were represented by the star-shaped symbol and the background points were represented by the circles. In Fig. 1.2, the two-dimensional data points $\mathbf{x} = (x, y)$ for each class were obtained by nonlinearly mapping the original Gaussian mixture data points $\mathbf{x}_0 = (x_0, y_0)$ in Fig. 1.1. All the data points in Fig. 1.2 were nonlinearly mapped by $\mathbf{x} = (x, y) = (x_0, x_0^2 + y_0)$. In the new data set the second component of each data point is nonlinearly related to its first component.

For both data sets, the contours generated by the kernel-based detectors are highly nonlinear and naturally following the dispersion of the data and thus successfully separating the two classes, as opposed to the linear contours obtained by the conventional detectors. Therefore, the kernel-based detectors clearly provided significantly improved discrimination over the conventional detectors for both the Gaussian mixture and nonlinearly mapped data. Among the kernel-based detectors, KMSD and KASD outperform KOSP and KSMF mainly because targets in KMSD and KASD are better represented by the associated target subspace than by a single spectral signature used in KOSP and KSMF. Note that the contour plots for MSD (Fig. 1.1(a) and Fig. 1.2 (a)) represent only the numerator of Eq. 1.5 because the denominator becomes unstable for the two-dimensional cases: i.e., the value inside the brackets $(\mathbf{I} - \mathbf{P}_{TB})$ becomes zero for the two-dimensional data.

1.7.2 Hyperspectral Images

In this section, HYDICE (HYperspectral Digital Imagery Collection Experiment) images from the Desert Radiance II data collection (DR-II) and Forest Radiance I data collection (FR-I) were used to compare detection performance between the kernel-based and conventional methods. The HYDICE imaging sensor generates 210 bands across the whole spectral range (0.4–2.5 $\mu$m) which includes the visible and short-wave infrared (SWIR) bands. But we use only 150 bands by discarding water absorption and low SNR bands; the spectral bands used are the 23rd–101st, 109th–136th, and 152nd–194th for the HYDICE images. The DR-II image includes 6 military targets along the road and the FR-I image includes total 14 targets along the tree line, as shown in the sample band images in Fig. 1.3. The detection performance of the DR-II and FR-I images was provided in both the qualitative and quantitative—the receiver operating characteristics (ROC) curves—forms. The spectral signatures of the desired target and undesired background signatures were directly

![Figure 1.3. Sample band images from (a) the DR-II image and (b) the FR-I image.](image-url)