

METODI DI RIVELAZIONE DI MATERIALI IN IMMAGINI IPERSPETTRALI RELATIVE A SCENARI COMPLESSI

Stefania MATTEOLI, Marco DIANI, Giovanni CORSINI

(*) Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Via Caruso 16, Pisa
Tel. +39 050 2217579 - Fax. +39 050 2217522
{stefania.matteoli, m.diani, g.corsini}@iet.unipi.it

Sommario

Questo lavoro analizza il problema dell'applicazione di metodi di *Matching Spettrale* ad immagini iperspettrali telerilevate relative a scenari complessi, caratterizzati da sfondo vario e fortemente disomogeneo. In tali situazioni, l'applicazione di metodi statistici, quali ad esempio l'*Adaptive Matched Filter* (AMF) e il *Fully Adaptive Detector* (FAD), tipicamente applicati utilizzando un Modello Gaussiano Locale (MGL), subisce una significativa degradazione delle prestazioni, dovuta all'inadeguatezza del MGL a modellare sfondi disomogenei, e all'effetto di contaminazione dovuto alla presenza di oggetti da rivelare molto vicini tra loro. Per operare in tali scenari, si ritiene più robusta la metodologia *Cluster Based* (CB), basata su una segmentazione dell'immagine e rivelazione effettuata entro ciascun *cluster*. Viene analizzata anche l'applicazione linea a linea dell'algoritmo deterministico *Constrained Energy Minimization* (CEM). In particolare, si intende valutare l'impatto sulle prestazioni della dimensione dei blocchi per il calcolo della matrice di correlazione. I metodi citati sono stati applicati ad un dato iperspettrale caratterizzato da forte disomogeneità, anche locale. Si è riscontrato che l'approccio CB non risente della degradazione delle prestazioni che invece affligge l'impiego del MGL. Per quanto riguarda il CEM, in tale scenario le migliori prestazioni sono ottenute stimando la matrice di correlazione su singola linea, o su blocchi di dimensione ridotta.

Abstract

This work presents an experimental analysis of *Spectral Matching* (SM) methods performed on a remotely sensed hyperspectral image of a very complex scenario, characterised by a strong local non-homogeneity in the background. In such a situation, the typical local applications of SM methods, such as the *Adaptive Matched Filter* (AMF) and the *Fully Adaptive Detector* (FAD), which relies upon a *Local Normal Model* (LNM), yields poor performance, due to the inadequateness of the LNM to represent local non-homogeneity. Moreover, the local application is also affected by contamination due to the presence of multiple objects to be detected in close proximity. In such scenarios, the *Cluster Based* (CB) methods seem to be more suitable, since they perform a cluster-conditioned detection, after having segmented the image. This allows us to avoid both the problems of local non-homogeneity and object contamination. The deterministic *Constrained Energy Minimisation* (CEM) algorithm is analysed as well. In its classical line by line implementation, the effect on the performance due to blocks size for correlation matrix estimation is investigated. Experimental results confirm the strength of CB approach to local non-homogeneity, compared to the performance degradation undergone by the LNM employment. As regards the CEM, a small blocks size, like a single line, provides the best performance.

Introduzione

L'elaborazione di immagini telerilevate permette di ottenere informazioni relative al territorio e all'ambiente, e trova dunque largo impiego in numerose applicazioni, quali, ad esempio, il monitoraggio del rischio ambientale, la rivelazione di materiali potenzialmente pericolosi e l'analisi della qualità delle acque superficiali. In particolare, negli ultimi anni, le ricadute applicative sono

divenute sempre più numerose, grazie allo sviluppo di sensori di ultima generazione, i *Sensori Iperspettrali* (SI). Tali sensori misurano il segnale ricevuto dalla scena osservata in centinaia di canali spettrali contigui, con una risoluzione spettrale elevatissima, anche dell'ordine del nanometro. Ciò conferisce all'analisi delle immagini da essi acquisite una potenziale estrema capacità di discriminare i differenti materiali presenti nella scena, sulla base della loro *firma spettrale* (FS), che caratterizza in modo univoco un dato materiale. In particolare, riferendosi alla porzione "riflettiva" dello spettro elettromagnetico (0.4-2.5 μm), per FS si intende la *Riflettanza Spettrale* (RS). L'enorme potenzialità del telerilevamento iperspettrale ha contribuito allo sviluppo della ricerca nell'ambito della rivelazione di materiali ed oggetti all'interno di una scena osservata, sulla base, appunto, di proprietà spettrali, piuttosto che spaziali. Ci riferiamo alle cosiddette tecniche di *Matching Spettrale* (MS), che si basano sulla conoscenza della FS del materiale di interesse per ricercare, entro l'immagine, quei *pixel* che presentano un elevato grado di correlazione con tale FS (Manolakis D., 2002, Manolakis D., 2005).

Poiché nelle applicazioni di MS tipicamente si ricercano oggetti *rari* (costituenti una minima frazione dello sfondo in cui sono immersi, come ad esempio un tetto di amianto in una vasta zona rurale), da un punto di vista statistico il problema è stato formulato in letteratura come un *test* ad ipotesi binaria, risolvibile come *Test a Rapporto di Verosimiglianza Generalizzato* (TRVG) (Kay S. M., 1998). I fattori chiave che permettono di classificare i vari algoritmi di MS sono essenzialmente il tipo di modello utilizzato per caratterizzare la variabilità spettrale (dello sfondo e dell'oggetto ricercato), la composizione del *pixel* in esame (detto *full-pixel* se puro, e *sub-pixel* se misto), ed il modello impiegato per descrivere i *pixel* misti (Manolakis D., 2002, Manolakis D., 2005). Poiché un esame esaustivo dei vari algoritmi esula dagli scopi di questo lavoro, in tale sede considereremo l'Adaptive Matched Filter (AMF) (Robey F.C. et al., 1992) ed il Fully Adaptive Detector (FAD) (Acito N. et al., 2005).

Nonostante l'approccio a TRVG sia il contesto naturale per la derivazione di algoritmi di MS, per completezza occorre precisare che alcuni algoritmi possono essere derivati seguendo un approccio di ottimizzazione (Manolakis D., 2002). Questo è, per esempio, il caso del *Constrained Energy Minimization* (CEM) (Chang C-I. et al., 2001).

L'obiettivo di questo lavoro è valutare l'impatto che uno scenario complesso fortemente disomogeneo ha sulle prestazioni di tali algoritmi, e quali siano le modalità più adatte di applicazione di tali metodi in situazioni di questo tipo.

Il presente lavoro è articolato come segue: verranno presentati brevemente gli algoritmi considerati, successivamente sarà esposta la problematica derivante dal loro impiego in scenari complessi e fortemente disomogenei, e verrà proposta una metodologia per la loro applicazione; di seguito, dopo una breve descrizione del dato iperspettrale utilizzato, verranno discussi i risultati sperimentali e le conclusioni.

Metodi di *Matching Spettrale* considerati

L'AMF è un algoritmo di rivelazione di oggetti *sub-pixel* che si basa su un modello dello sfondo di tipo non strutturato (Manolakis D., 2002, Manolakis D., 2005).

Di seguito si riportano le due ipotesi e la formula del rivelatore, rimandando per ulteriori dettagli alla letteratura (Robey F.C. et al., 1992, Manolakis D., 2002, Manolakis D., 2005):

$$\begin{matrix} \underline{x} | H_0 \sim N(\underline{\mu}, \underline{C}) \\ \underline{x} | H_1 \sim N(\underline{a} + \underline{\mu}, \underline{C}) \end{matrix} \quad \left(\text{AMF}(\underline{x}) = \frac{\left[(\underline{s} \ \underline{\mu})' \underline{\hat{C}}^{-1} (\underline{x} \ \underline{\mu}) \right]^2}{(\underline{s} \ \underline{\mu})' \underline{\hat{C}}^{-1} (\underline{s} \ \underline{\mu})} \right) \begin{matrix} > \eta \\ < \eta \end{matrix} \quad \% [1]$$

In [1], $\underline{x} = [x_1, x_2, \dots, x_L]^T$ rappresenta il vettore aleatorio multidimensionale usato per modellare il *pixel* in esame, relativo ad un'immagine iperspettrale ad L bande; la sua realizzazione è indicata con $\underline{x} = [x_1, x_2, \dots, x_L]^T$. La FS dell'oggetto ricercato è indicata con $\underline{s} = [s_1, s_2, \dots, s_L]^T$. Con \underline{C} e $\underline{\eta}$ si indicano la matrice di covarianza e il vettore medio dello sfondo e con $\underline{\hat{C}}$ e $\underline{\hat{\eta}}$ le rispettive stime a massima verosimiglianza, delle quali si discuterà in maggior dettaglio nella sezione successiva.

Il FAD è stato derivato invece nel caso *full-pixel* utilizzando un modello a sostituzione dello sfondo per le due ipotesi (Acito N. et al., 2005). Se ne riportano ipotesi e formula, secondo la medesima notazione:

$$\begin{cases} \underline{\mathbf{X}} | H_0 \in N(\underline{\boldsymbol{\eta}}, \underline{\mathbf{C}}) \\ \underline{\mathbf{X}} | H_1 \in N(\underline{\mathbf{s}}, \underline{\mathbf{C}}) \end{cases} \Rightarrow \text{FAD}(\underline{\mathbf{x}}) = \frac{(\underline{\mathbf{s}} - \underline{\hat{\boldsymbol{\eta}}})^t \cdot \hat{\underline{\mathbf{C}}}^{-1} \cdot (\underline{\mathbf{x}} - \underline{\hat{\boldsymbol{\eta}}})}{\sqrt{(\underline{\mathbf{s}} - \underline{\hat{\boldsymbol{\eta}}})^t \cdot \hat{\underline{\mathbf{C}}}^{-1} \cdot (\underline{\mathbf{s}} - \underline{\hat{\boldsymbol{\eta}})}}} \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \lambda \quad [2]$$

Per quanto riguarda il CEM, questo è stato ottenuto realizzando il filtro *Finite Impulse Response* (FIR) \underline{h} che minimizza l'energia in uscita al filtro stesso, con il vincolo di ottenere una risposta unitaria alla FS di interesse (Chang C-I. et al., 2001). La derivazione del CEM è sintetizzata di seguito:

$$\begin{cases} \min_{\underline{h}} \{ \underline{h}^t \cdot \hat{\underline{\mathbf{R}}}_x \cdot \underline{h} \} \\ \underline{h}^t \cdot \underline{\mathbf{s}} = 1 \end{cases} \Rightarrow \text{CEM}(\underline{\mathbf{x}}) = \frac{\underline{\mathbf{s}}^t \cdot \hat{\underline{\mathbf{R}}}_x^{-1} \cdot \underline{\mathbf{x}}}{\underline{\mathbf{s}}^t \cdot \hat{\underline{\mathbf{R}}}_x^{-1} \cdot \underline{\mathbf{s}}} \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \lambda \quad [3]$$

dove con $\hat{\underline{\mathbf{R}}}_x$ si è indicata la matrice di correlazione di $\underline{\mathbf{X}}$ stimata a massima verosimiglianza.

Problematica e metodologia

Da un punto di vista pratico, per realizzare gli algoritmi enunciati nella sezione precedente, occorre determinare le stime a massima verosimiglianza del vettore medio $\underline{\hat{\boldsymbol{\eta}}}$ e della matrice di covarianza $\hat{\underline{\mathbf{C}}}$ dello sfondo, e della matrice di correlazione $\hat{\underline{\mathbf{R}}}_x$, utilizzando parte dei *pixel* dell'immagine. Per quanto riguarda $\underline{\hat{\boldsymbol{\eta}}}$ e $\hat{\underline{\mathbf{C}}}$, il loro calcolo differisce a seconda di come si intende modellare lo sfondo. Uno degli approcci più diffusi è quello di assumere un *Modello Gaussiano Locale* (MGL), ampiamente utilizzato per la sua trattabilità matematica. Tale assunzione può essere ragionevolmente intrapresa nell'ambito di un modello *non stazionario*: l'immagine iperspettrale può essere cioè modellata come un processo Gaussiano non stazionario con un vettore medio rapidamente variabile spazialmente ed una matrice di covarianza meno rapidamente variabile (Hunt B. R., Cannon T. M., 1976). Ciò significa che $\underline{\hat{\boldsymbol{\eta}}}$ e $\hat{\underline{\mathbf{C}}}$ possono essere stimati ragionevolmente mediante l'impiego di una finestra mobile (Margalit A. et al., 1985, Chen J. Y., Reed I. S., 1987) che scorre lungo l'immagine. Da un punto di vista pratico, per evitare stime polarizzate a causa della possibile presenza dell'oggetto di interesse nei *pixel* intorno a quello in esame, occorre utilizzare una finestra di guardia (Acito N. et al., 2005) che escluda tali *pixel*. Inoltre, la dimensione della finestra per la stima di $\hat{\underline{\mathbf{C}}}$ deve essere adeguatamente ampia per consentire una stima accurata di $\hat{\underline{\mathbf{C}}}$ (almeno $10 L$, (Richards J. A., Jia X., 1993)), ma adeguatamente ridotta per catturare sfondo di tipo omogeneo.

Tuttavia, il MGL non riesce a modellare adeguatamente lo sfondo in presenza di scenari complessi, nei quali si ha disomogeneità locale causata dalla presenza di vari oggetti/materiali nell'immediata vicinanza del *pixel* in esame. Inoltre, in certi scenari, è possibile che gli oggetti di interesse si trovino molto vicini l'uno all'altro, determinando una "contaminazione" nelle stime anche utilizzando la finestra di guardia. In tali casi, per evitare una significativa degradazione delle prestazioni del rivelatore impiegato, occorre adottare un modello più adeguato per la rappresentazione dello sfondo, che non sia affetto dalla disomogeneità locale intrinseca all'immagine analizzata. Il *Modello a Mistura di Gaussiane* (MMG) modella lo sfondo come costituito da più classi spettralmente omogenee, per ciascuna delle quali può essere assunta la gaussianità (condizionatamente alla classe). L'approccio tipico in questo caso è quello cosiddetto *Cluster Based* (CB), comunemente adottato per la rivelazione di anomalie spettrali (Ashton E. A., 1998, Carlotto M. J., 2005), nel quale l'immagine viene segmentata tramite l'impiego di un classificatore non supervisionato, come ad esempio lo *Stochastic Expectation Maximization* (SEM), e la rivelazione viene eseguita condizionatamente al *cluster* di appartenenza del *pixel* in esame.

Ovviamente occorre che i *cluster* formati consistano di un numero di campioni sufficiente per stimare le statistiche sul *cluster* ed invertire \hat{C} . Inoltre, non si deve sovrastimare il numero di *cluster*, in modo da evitare che gli oggetti da rivelare formino *cluster* a sé stanti. Ci si aspetta che l'approccio CB riesca ad evitare la degradazione delle prestazioni subita impiegando il MGL nel caso di scenari complessi con disomogeneità a livello locale.

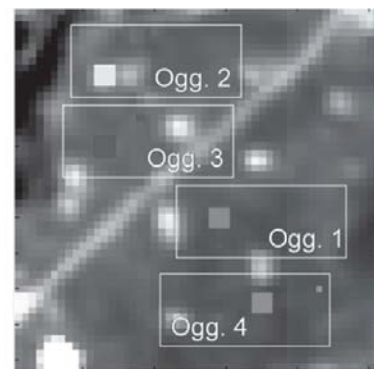
Per quanto riguarda il CEM, ci si attiene alla formulazione classica (Chang C-I. et al., 2001), con implementazione in tempo reale elaborando l'immagine linea a linea. Relativamente alla stima di \hat{R}_x , questa viene aggiornata linea dopo linea. In particolare, tale stima può essere effettuata *ex-novo* per ogni linea acquisita, non tenendo memoria delle stime effettuate alle linee precedenti, oppure può essere determinata aggiornando la stima calcolata alle linee precedenti. Si può poi introdurre un cosiddetto *forgetting-factor*, che permette di non tenere in memoria le stime effettuate in corrispondenza di linee ormai lontane. Variando tale fattore, si varia di fatto la dimensione del blocco di linee che si utilizza per la stima di \hat{R}_x , che può variare da essere una linea sola, fino ad arrivare all'intera dimensione dell'immagine. Chiaramente, la dimensione del blocco utilizzato ha un impatto sulle prestazioni del rivelatore, in dipendenza anche dal tipo di scenario (disomogeneità, ecc.). Utilizzando infatti per la stima, linea dopo linea, sempre più campioni, aumenta il rango della matrice, ma si catturano porzioni sempre più vaste di sfondo, col rischio di incorrere in disomogeneità. In presenza di scenari fortemente disomogenei, ci si aspettano migliori prestazioni con una stima effettuata su singola linea, o al limite su blocchi di dimensioni ridotte.

Descrizione dei dati utilizzati

L'immagine iperspettrale utilizzata in questa sede è stata acquisita da un sensore iperspettrale operante nell'intervallo spettrale 0.4-2.5 μm con 126 bande e risoluzione spettrale dell'ordine dei 16 nm. L'immagine consta di 280x800 *pixel*. La scena acquisita è relativa ad uno scenario rurale piuttosto complesso, comprendente un villaggio, e dunque edifici e veicoli, oltre a vegetazione, erba, strade ed un fiume. La scena è riportata in Figura 1 (a). Nella scena sono stati posti dei teli di diverso materiale, colorazione, e dimensione, le cui caratteristiche sono riassunte in Tabella 1. La maschera di verità a terra relativa a tali oggetti è rappresentata in Figura 1 (b).



(a)



(b)

Figura 1. (a) Rappresentazione della scena acquisita; (b) Mappa di verità a terra relativa agli oggetti da rivelare.

Nome	Colorazione	Unità	Dimensioni
Ogg.1	Rosso	1	3m x 3m
Ogg.2	Giallo	1	3m x 3m
Ogg.3	Blu	2	2m x 2m, 1m x 1m
Ogg.4	Rosso	2	2m x 2m, 1m x 1m

Tabella 1. Caratteristiche degli oggetti da rivelare

L'intervallo di campionamento spaziale a terra è dell'ordine dei 3m, dunque, come si evince dalla Tabella 1, ci si trova in una situazione prevalentemente *sub-pixel* (al più un *pixel* pieno nei casi migliori). Dalle figure 1 (a) e (b) è evidente come lo sfondo sia fortemente disomogeneo, anche localmente in prossimità degli oggetti di interesse; inoltre, si nota la presenza di oggetti, anche piuttosto estesi, con colorazione simile a quella dei teli da rivelare. È inoltre evidente come, la vicinanza dei teli uno all'altro renda effettivamente complesso l'utilizzo di una finestra mobile per la stima dei parametri dello sfondo.

Risultati sperimentali e discussione

In questa sezione sono riportati i risultati sperimentali ottenuti applicando gli algoritmi enunciati con la metodologia proposta nella sezione precedente.

Poiché gli oggetti da rivelare constano di un numero esiguo di *pixel*, le prestazioni saranno comparate, oggetto a oggetto, in termini del numero di falsi allarmi (#FA) ottenuti per rivelare almeno un *pixel* dell'oggetto stesso.

Per l'applicazione di AMF e FAD in locale, la dimensione della finestra mobile è stata presa 15 x 15 e 9 x 9, ritenuto un buon compromesso per evitare il più possibile (ma non totalmente) la contaminazione da oggetti vicini e catturare un numero sufficiente di campioni per la stima delle statistiche dello sfondo. Per l'approccio CB, è stato effettuato un *clustering* dell'immagine su 10 *cluster*. La Figura 2 (a) riporta in grafico il #FA ottenuto nei vari casi per la rivelazione dei differenti oggetti. È evidente come la degradazione delle prestazioni chiaramente osservabile nell'applicazione locale venga completamente superata dai metodi CB, per i quali si ottiene una rivelazione senza falsi allarmi per tre su quattro oggetti.

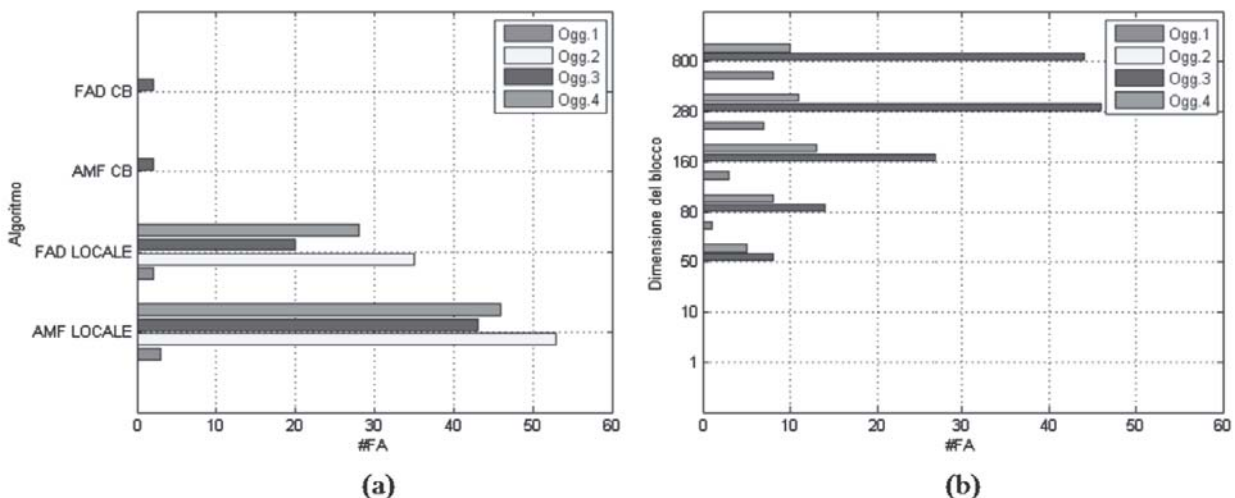


Figura 2. Numero di falsi allarmi (#FA) ottenuti sui quattro tipi di oggetto: (a) applicazione locale e CB di AMF e FAD; (b) applicazione del CEM al variare della dimensione del blocco.

Per quanto riguarda il CEM, si è fatta variare la dimensione del blocco di linee secondo i valori {1, 10, 50, 80, 160, 280, 800}. Dall'analisi della Figura 2 (b), si riscontra che il #FA ha un andamento

che tende a diminuire man mano che diminuisce la dimensione del blocco. Tale andamento è riscontrabile nella rivelazione dei differenti oggetti. In particolare, si ottiene che per blocchi di 10 linee e di una singola linea tutti gli oggetti vengono rivelati senza falsi allarmi.

Conclusioni

In questo lavoro è stata analizzata l'applicazione di algoritmi di MS a scenari complessi con sfondo fortemente disomogeneo. Sono stati considerati AMF e FAD nell'ambito degli algoritmi "TRVG-derivati", e il CEM per quanto riguarda i metodi deterministici. Per AMF e FAD si è riscontrato che l'approccio CB è più robusto, rispetto al MGL, alla disomogeneità dello sfondo. Inoltre, l'approccio CB è applicabile anche in presenza di oggetti da rivelare molto vicini tra loro, e dunque non risente di problemi di contaminazione. Sono state successivamente analizzate le prestazioni ottenute col CEM variando la dimensione dei blocchi per la stima della matrice di correlazione. Nello scenario analizzato, fortemente vario e disomogeneo, una stima su singola linea, o su blocchi di dimensioni ridotte, produce le migliori prestazioni.

Ringraziamenti

Si ringrazia Dr. J. Kerekes (Rochester Institute of Technology) per aver fornito i dati utilizzati in questa analisi.

Bibliografia

- Acito N. et al. (2005), "Adaptive Detection algorithm for full-pixel targets in hyperspectral images", *IEE Proceedings-Vision, Image, and Signal Processing*, 152(6): 731-740.
- Ashton E. A. (1998), "Detection of Subpixel Anomalies in Multispectral Infrared Imagery Using an Adaptive Bayesian Classifier", *IEEE Transactions on Geoscience and Remote Sensing*, 36(2): 506-517.
- Carlotto M. J. (2005), "A Cluster-Based Approach for Detecting Man-Made Objects and Changes in Imagery", *IEEE Transactions on Geoscience and Remote Sensing*, 43(2): 374-387.
- Chang C-I. et al. (2001), "Real-Time Processing Algorithms for Target Detection and Classification in Hyperspectral Imagery", *IEEE Transactions on Geoscience and Remote Sensing*, 39(4): 760-768.
- Chen J. Y., Reed I. S. (1987), "A Detection Algorithm for Optical targets in Clutter", *IEEE Transactions on Aerospace and Electronic Systems*, 23(1) 46-59.
- Hunt B. R., Cannon T. M. (1976), "Nonstationary Assumptions of Gaussian Models of Images", *IEEE Transactions on Systems, Man, and Cybernetics*, 6(6), 876-882.
- Kay S. M. (1998), *Fundamentals of statistical signal processing: Detection Theory*, Prentice Hall, New Jersey.
- Manolakis D. (2005), "Taxonomy of detection algorithms for hyperspectral imaging applications", *Optical Engineering*, 44(6), 1-11.
- Manolakis D., Shaw G. (2002), "Detection Algorithms for Hyperspectral Imaging Applications", *IEEE Signal Processing Magazine*, 29-43.
- Margalit A. et al. (1985), "Adaptive Optical Target Detection Using Correlated Images", *IEEE Transactions on Aerospace and Electronic Systems*, 21(3): 394-405.
- Richards J. A., Jia X., *Remote Sensing Digital Image Processing*, Springer-Verlag, Germany.
- Robey F. C. et al. (1992), "A CFAR adaptive matched filter detector", *IEEE Transactions on Aerospace and Electronic Systems*, 28(1): 208-216.