Discrimination and Identification for Subpixel Targets in Hyperspectral Imagery

Chein-I Chang[1] Weimin Liu[1] Chein-Chi Chang[2]
[1]Remote Sensing Signal and Image Processing Laboratory
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County, Baltimore, MD 21250
[2]Department of Civil and Environmental Engineering
University of Maryland, Baltimore County, Baltimore, MD 21250

Abstract

Spectral measures have been used in material identification and discrimination. They are effective if the spectral signatures are calibrated and not contaminated. However, it may not be true in many real applications, specifically, for mixed pixels and subpixel targets. This paper investigates the issue of discrimination and identification for subpixel targets and further develops sample spectral covariance/correlation matrix-based hyperspectral measures to account for spectral variability within subpixel targets. Two types of measures are of interest and studied, Mahalanobis distance-based hyperspectral measures and matched filter-based hyperspectral measures. In order to substantiate the proposed measures, a real data-based comparative analysis is conducted and compared to two spectral similarity measures, spectral angle mapper (SAM) and spectral information divergence (SID) for performance evaluation. The experiments show that both Mahalanobis distance-based hyperspectral measures and matched filter-based hyperspectral measures work very effectively and outperformed the SAM and the SID in discrimination and identification for subpixel targets.

I. INTRODUCTION

Spectral similarity provides an important feature in material identification, discrimination, detection and classification. Many pure pixel-based spatial classical similarity measures such as spectral angle mapper (SAM) [1], and Euclidean distance (ED) [2] have been used for this purpose. Recently, a new stochastic measure, called spectral information divergence (SID) [2-3] was particularly developed for hyperspectral imagery. All of these measures are designed to be performed on a single pixel vector basis to capture spectral similarity between two pixel vectors. In this case, they are effective if the spectral signatures are calibrated and not contaminated, specifically, for full pixel. However, this idealistic case is generally not true in many real hyperspectral images where many factors may cause contamination of spectral signatures to be identified. One scenario is mixed pixel discrimination and identification where the spectral signature of a pixel vector is mixed with a set of signatures resident in the pixel. Another is subpixel target discrimination and identification where the target to be identified is embedded in a single pixel vector and its spectral signature is clearly mixed with other signatures that are also present in the pixel vector. In either case, pixel-based spectral measures may be very likely to fail and sometimes may even identify wrong targets. This paper investigates the issue of discrimination and identification for subpixel targets, and provides evidence to demonstrate that such examples indeed occur in real hyperspectral imagery where some commonly used spectral measures not only fail in discrimination and identification of subpixel targets, but also identify wrong targets. In order to address this problem, this paper further develops new measures for subpixel target identification and discrimination. Unlike pixel-based spectral measures as described above, these measures take advantage of the sample spectral correlation to account for spectral variability of subpixel targets. Two types of sample spectral correlation-based spectral measures previously developed as target discrimination measures for anomaly classification in hyperspectral imagery [4] will be studied in this paper. They are Mahalanobis distance-based and matched filter-based hyperspectral measures which include the sample spectral covariance/correlation matrix to capture spectral signatures of subpixel targets more effectively. Therefore, these measures can be considered as second-order (2^{nd}-order) spectral measures compared to pixel-based spectral measures that do not include the sample covariance/correlation matrix and can be thought of as first-order (1^{st}-order) spectral measures. In order to demonstrate the effectiveness of the proposed 2^{nd}-order spectral measures a comparative study on performance analysis between the 1^{st}-order and 2^{nd}-order spectral measures is conducted via real hyperspectral imagery. As experiments show, the 1^{st}-order pixel-based spectral measures can only be effective in discrimination and identification of pure pixels, not necessarily for subpixel targets. To the contrary, the 2^{nd}-order spectral measures perform very impressively in subpixel discrimination and identification.

II. PIXEL-BASED SPECTRAL MEASURES

In this section, two spectral discrimination measures will be presented and used to measure similarity between any two-pixel vectors. Assume that $\mathbf{s}_i = (s_{i1}, s_{i2}, \cdots, s_{iL})^T$ and $\mathbf{s}_j = (s_{j1}, s_{j2}, \cdots, s_{jL})^T$ are two spectral signatures.

A. Spectral Angle Mapper (SAM) [1]

The SAM measures spectral similarity by finding the angle between the spectral signatures $\mathbf{s}_i$ and $\mathbf{s}_j$

0-7803-8554-3/04/$20.00 ©2004 IEEE.

$$\mathrm{SAM}(\mathbf{s}_i,\mathbf{s}_j)=\cos^{-1}\left(\frac{\langle \mathbf{s}_i,\mathbf{s}_j\rangle}{\|\mathbf{s}_i\|\|\mathbf{s}_j\|}\right) \quad (1)$$

where $\langle \mathbf{s}_i,\mathbf{s}_j\rangle=\sum_{l=1}^{L}s_{il}s_{jl}$, $\|\mathbf{s}_i\|=\left(\sum_{l=1}^{L}s_{il}^2\right)^{1/2}$ and $\|\mathbf{s}_j\|=\left(\sum_{l=1}^{L}s_{jl}^2\right)^{1/2}$.

B. Spectral Information Divergence (SID) [3]

Let $\mathbf{p}=(p_1,p_2,\ldots,p_L)^T$ and $\mathbf{q}=(q_1,q_2,\ldots,q_L)^T$ the two probability mass functions generated by $\mathbf{s}_i$ and $\mathbf{s}_j$ respectively with $p_l=s_{il}/\sum_{l=1}^{L}s_{il}$ and $q_l=s_{jl}/\sum_{l=1}^{L}s_{jl}$. So, the self-information provided by $\mathbf{s}_i$ and $\mathbf{s}_j$ for band l is defined by $I_l(\mathbf{s}_i)=-\log p_l$ and $I_l(\mathbf{s}_j)=-\log q_l$ respectively and the entropy of $\mathbf{s}_j$ relative to $\mathbf{s}_i$ is the defined by

$$D(\mathbf{s}_i\|\mathbf{s}_j)=\sum_{l=1}^{L}p_l\log(p_l/q_l) \quad (2)$$

which is also known as Kullback-Leibler information measure, directed divergence or cross entropy [5]. Similarly, we can also define the average discrepancy in the self-information of $\mathbf{s}_i$ relative to the self-information of $\mathbf{s}_j$ by

$$D(\mathbf{s}_j\|\mathbf{s}_i)=\sum_{l=1}^{L}q_l\log(q_l/p_l) \quad (3)$$

Summing Eqs. (2) and (3) yields Spectral Information Diveregence (SID) [2,3]

$$\mathrm{SID}(\mathbf{s}_i,\mathbf{s}_j)=D(\mathbf{s}_i\|\mathbf{s}_j)+D(\mathbf{s}_j\|\mathbf{s}_i), \quad (4)$$

which can be used to measure the discrepancy between two pixel vectors $\mathbf{s}_i$ and $\mathbf{s}_j$ in terms of their corresponding probability mass functions, $\mathbf{p}$ and $\mathbf{q}$.

III. HYPERSPECTRAL MEASURES FOR SUBPIXEL TARGETS

The SAM and SID measure the spectral similarity between a pair of two pixel vectors on a single pixel basis. In other words, the pixel's similarity is measured only based on the information provided by a pixel vector. So, if a material to be measured does not occupy a full pixel, the spectral characteristics of the pixel are not necessarily to be represented well the spectral properties of the material. This often occurs in real applications when a material is either mixed with other signatures such as background or embedded in a single pixel as a subpixel target. In both cases, using full-pixel spectral information measure material similarity is generally not effective as will be demonstrated in the experiments in the following section. In order to resolve this difficulty, we need to develop new measures that can measure spectral characteristics up to subpixel level. As noted, most measures including SAM and SID used for spectral similarity do not take into account the spectral correlation among pixels, which is very useful and crucial information in identification or discrimination of materials in real images. One scenario was demonstrated in [4] where the four target discrimination measures were developed for anomaly classification. Interestingly, their applications of these measures to subpixel target discrimination and identification was not explored and further studied in [4]. This section presents these four measures and illustrates how they can be used for identification and discrimination of subpixel targets. Generally speaking, these four measures can be categorized into two classes, Mahalanobis-based measures and matched filter-based measures.

Two Mahalanobis-based distance measures are covariance-Mahalanobis distance (CMD) and correlation-Mahalanobis distance (RMD)

$$\mathrm{CMD}(\mathbf{s}_i,\mathbf{s}_j)=(\mathbf{s}_i-\mathbf{s}_j)^T\mathbf{K}_{L\times L}^{-1}(\mathbf{s}_i-\mathbf{s}_j) \quad (5)$$

$$\mathrm{RMD}(\mathbf{s}_i,\mathbf{s}_j)=(\mathbf{s}_i-\mathbf{s}_j)^T\mathbf{R}_{L\times L}^{-1}(\mathbf{s}_i-\mathbf{s}_j) \quad (6)$$

where $\mathbf{K}_{L\times L}$ and $\mathbf{R}_{L\times L}$ are sample covariance and correlation matrices respectively.

Two matched filter-based distance measures are covariance matched filter-based distance (CMFD) and correlation matched filter-based distance (RMFD) are defined by

$$\mathrm{CMFD}(\mathbf{s}_i,\mathbf{s}_j)=(\mathbf{s}_i-\boldsymbol{\mu})^T\mathbf{K}_{L\times L}^{-1}(\mathbf{s}_j-\boldsymbol{\mu}) \quad (7)$$

$$\mathrm{RMFD}(\mathbf{s}_i,\mathbf{s}_j)=\mathbf{s}_i^T\mathbf{R}_{L\times L}^{-1}\mathbf{s}_j \quad (8)$$

where $\boldsymbol{\mu}$ is the sample mean of the image.

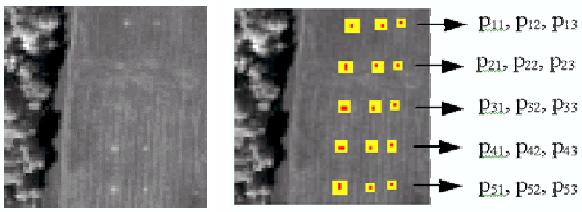
As we can see from Eqs. (5-8), all the measures involve calculation of either sample spectral covariance matrix or sample spectral correlation matrix which are not included in any pure pixel-based spectral similarity measure. They can be referred to as 2nd-order spectral measures as opposed to SAM and SID which can be considered as the 1st-order spectral measures. As evidence shown in the following experiments, it is the sample spectral correlation in 2nd-order spectral measures make significant improvement in measuring spectral similarity among subpixel targets.

IV. EXPERIMENTS

HYperspectral Digital Image Collection Experiment (HYDICE) image data in [3] were used for experiments to demonstrate the superior performance of the 2nd-order spectral measures over the 1st-order spectral measures in subpixel discrimination and identification.

The real data to be used for experiments are a HYDICE image scene shown in Fig. 1(a), which has a size of 64×64 pixel vectors with 15 panels in the scene with the ground truth map provided in Fig. 1(b). It was acquired by 210 spectral bands with a spectral coverage from 0.4μm to 2.5 μm. Low signal/high noise bands: bands 1-3 and bands 202-210; and water vapor absorption bands: bands 101-112 and bands 137-153 were removed. So, a total of 169 bands were used. The spatial resolution is 1.56m and spectral resolution is 10nm. Within the scene in Fig. 1(a) there is a large grass

field background, and a forest on the left edge. Each element in this matrix is a square panel and denoted by p_{ij} with rows indexed by i and columns indexed by $j = 1,2,3$. For each row $i = 1,2,\text{?},5$, there are three panels p_{i1}, p_{i2}, p_{i3}, painted by the same material but with three different sizes. For each column $j = 1,2,3$, the 5 panels p_{1j}, p_{2j}, p_{3j}, p_{4j}, p_{5j} have the same size but with five different materials. The sizes of the panels in the first, second and third columns are $3\text{m} \times 3\text{m}$, $2\text{m} \times 2\text{m}$ and $1\text{m} \times 1\text{m}$ respectively. Since the size of the panels in the third column is $1\text{m} \times 1\text{m}$, they cannot be seen visually from Fig. 1(a) due to the fact that its size is less than the 1.56m pixel resolution. Fig. 1(b) shows the precise spatial locations of these 15 panels where red pixels (R pixels) are the panel center pixels and the pixels in yellow (Y pixels) are panel pixels mixed with the background.



(a) image scene (b) ground truth map of 15 panels

Figure 1. HYDICE image

The 1.56m-spatial resolution of the image scene suggests that most of the 15 panels are one pixel in size except that p_{21}, p_{31}, p_{41}, p_{51} which are two-pixel panels. Fig. 2 plots the 5 panel spectral signatures P_i for $i = 1,2,\text{?},5$ obtained by averaging R pixels in Fig. 1(b), where the i-th panel signature, denoted by P_i was generated by averaging the red panel center pixels in row i.

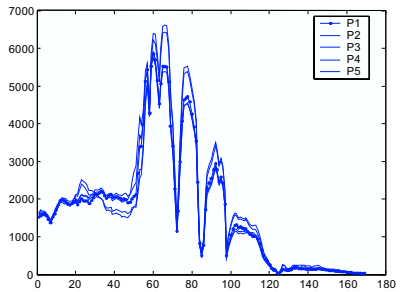


Figure 2. Five panel signatures P_i for $i = 1,2,\text{?},5$

These panel signatures represent the target knowledge of the panels in each row and will be used for identification of the pixels in the 15 panels. Tables 1-3 tabulate the identification results of the 1st-order spectral measures, SAM, SID, the Mahalanobis distance-based measures CMD, RMD, and matched filter-based measures, CMFD, RCMD respectively. In these tables, a correct identification is highlighted by shade, while a cross indicates a wrong identification. As we can see from these tables, SAM and SID made 5 errors, of which four errors were made on all the subpixel targets except p_{23} plus one error on pixel p_{41}. In contrast, all the 2nd-order spectral similarity measures performed significantly better than did the 1st-order spectral measures. All of these four measures identified the five subpixel targets, p_{13}, p_{23}, p_{33}, p_{43}, p_{53} correctly. This experiment clearly demonstrated that the 2nd-order spectral measures worked very well to identify and discriminate subpixel targets.

V. CONCLUSIONS

This paper investigates the issue of subpixel discrimination and identification. Over the past years, spectral measures such as SAM and Euclidean distance have been used for this purpose. Unfortunately, they are generally used for pure pixel-based spatial-based classical image processing and may not be effective when targets are embedded in a single pixel. In this case, discrimination and identification of subpixel targets can be only performed at subpixel level. Therefore, in this paper, one of major contributions of this paper is to provide evidence that the commonly used spatial-based 1st-order spectral measures are not generally effective in subpxiel discrimination and identification. Another is to develop 2nd-order hyperspectral measures that improve the 1st-order spectral measures and also work effectively for subpixel discrimination and identification. A third contribution is to offer new insights into subpixel detection and future research in anomaly detection and mixed pixel discrimination and identification.

References

[1] R.A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, 2nd ed., Academic Press, 1997.

[2] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral detection and Classification*, Kluwer Academic/Plenum Publishers, 2003.

[3] C.-I Chang, "An information theoretic-based approach to spectral variability, similarity and discriminability for hyperspectral image analysis," *IEEE Trans. on Information Theory*, vol. 46, no. 5, pp. 1927-1932, August 2000.

[4] C.-I Chang and S.-S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no. 6, pp. 1314-1325, June 2002.

[5] T. Cover and J. Thomas, Elements of Information Theory, John Wiley & Sons, 1991.

Table 1. Discrimination and identification of the pixels in the 15 panels using SAM and SID (each made 5 errors)

	SAM					SID				
	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5
P_{11}	0.0531	0.0873	0.0979	0.1008	0.105	0.0059	0.0138	0.0173	0.0123	0.0138
P_{12}	0.0079	0.0411	0.0658	0.116	0.1258	0.0002	0.0022	0.0052	0.0173	0.0234
P_{13}	0.0579	~~0.0453~~	0.074	0.1516	0.164	0.0068	~~0.0039~~	0.0069	0.0338	0.0437
P_{21}	0.045	0.016	0.0387	0.1495	0.157	0.0026	0.0005	0.0024	0.0263	0.0323
P_{22}	0.0365	0.0163	0.0462	0.1375	0.1461	0.0022	0.0004	0.0025	0.0233	0.03
P_{23}	0.066	0.0338	0.0647	0.1643	0.1752	0.0066	0.0019	0.005	0.0353	0.0445
P_{31}	0.081	0.0818	0.054	0.1561	0.1571	0.0083	0.0085	0.0041	0.0267	0.0301
P_{32}	0.0849	0.0477	0.0469	0.1827	0.1915	0.0085	0.0025	0.0025	0.0402	0.0483
P_{33}	0.1066	~~0.0684~~	0.0768	0.2028	0.2131	0.0152	~~0.0064~~	0.0072	0.053	0.0635
P_{41}	0.1538	0.1897	0.2036	0.0464	~~0.0408~~	0.0319	0.0474	0.054	0.0035	~~0.0026~~
P_{42}	0.1174	0.1505	0.1676	0.01	0.0252	0.0176	0.0286	0.0352	0.0002	0.0014
P_{43}	0.0777	~~0.0646~~	0.0963	0.1429	0.1568	0.0069	~~0.0039~~	0.0092	0.0275	0.0366
P_{51}	0.1387	0.1717	0.1852	0.0369	0.0177	0.0283	0.0418	0.048	0.0035	0.0006
P_{52}	0.1526	0.1863	0.1978	0.0509	0.0326	0.0347	0.05	0.0556	0.0058	0.0019
P_{53}	0.0748	~~0.0568~~	0.0889	0.147	0.1597	0.0067	0.0031	0.008	0.0289	0.0375

Table 2. Discrimination and identification of the pixels in the 15 panels using CMD and RMD (each made 0 error)

	CMD					RMD				
	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5
P_{11}	380.12	2007.10	1996.90	2021.00	1963.30	379.46	2007.00	1992.30	2021.00	1962.40
P_{12}	190.34	1069.60	994.32	1091.70	1064.10	190.17	1068.60	993.47	1090.70	1064.10
P_{13}	355.00	609.21	754.57	704.18	790.89	354.84	608.26	753.71	703.20	790.83
P_{21}	1241.50	126.84	1098.50	1229.70	1255.40	1236.00	123.71	1085.10	1226.60	1249.20
P_{22}	1205.80	178.94	1155.20	1214.20	1215.30	1204.00	175.24	1155.20	1210.40	1213.90
P_{23}	792.64	342.14	699.94	782.54	860.23	792.37	340.97	699.27	781.33	860.10
P_{31}	1365.30	1105.60	178.51	1439.80	1393.80	1363.00	1101.30	178.49	1435.50	1392.00
P_{32}	949.55	851.35	170.99	991.63	1010.60	947.27	847.04	170.96	987.25	1008.70
P_{33}	750.86	700.03	249.64	812.91	815.21	750.68	700.01	246.57	812.88	814.88
P_{41}	1520.30	1461.70	1590.50	249.33	1420.10	1520.10	1461.70	1587.40	249.30	1419.80
P_{42}	1198.10	1194.00	1262.20	168.52	1133.90	1198.10	1193.60	1260.50	168.16	1133.90
P_{43}	753.09	743.25	781.89	443.25	686.90	752.58	743.23	777.72	443.23	686.16
P_{51}	1250.50	1205.80	1230.40	1160.20	159.00	1250.50	1205.60	1228.20	1160.10	158.90
P_{52}	1260.80	1325.70	1285.70	1160.40	222.44	1260.80	1325.20	1284.30	1159.80	222.44
P_{53}	737.39	661.75	717.23	677.86	427.13	736.09	658.82	717.19	674.88	426.14

Table 3. Discrimination and identification of the pixels in the 15 panels using CMFD and RMFD (each made 0 error)

	CMFD					RMFD				
	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5
P_{11}	819.664	-24.318	9.710	11.996	32.605	818.8952	-23.8813	9.6279	11.9344	31.7162
P_{12}	441.602	-28.520	38.061	3.712	9.233	443.6709	-24.5505	39.1888	7.2021	10.9981
P_{13}	189.191	31.583	-12.138	27.393	-24.229	186.3707	30.6611	-15.8956	25.9916	-27.3525
P_{21}	-8.796	518.019	61.126	9.891	-11.215	-9.5839	517.5706	63.053	8.92	-11.8947
P_{22}	-5.861	477.079	17.899	2.746	-6.096	-1.6461	483.7324	19.9137	8.9342	-2.3267
P_{23}	-7.331	187.410	37.470	10.507	-36.606	-7.1584	189.5439	36.559	12.1625	-36.753
P_{31}	3.561	102.896	595.386	-20.937	-6.194	3.6542	105.5097	593.0835	-18.7855	-6.5694
P_{32}	13.809	32.404	401.543	5.561	-12.183	12.1557	33.2765	397.4838	5.9709	-14.3054
P_{33}	19.215	14.125	268.277	0.979	-8.435	16.7792	13.1188	266.0042	-0.52	-11.0499
P_{41}	18.836	17.645	-17.778	667.120	73.450	21.017	21.2502	-15.4252	670.2322	75.4524
P_{42}	35.233	6.778	1.626	562.792	71.856	37.3921	10.6177	3.3581	566.1465	73.7695
P_{43}	10.063	-15.525	-5.886	177.771	47.679	8.9613	-15.3621	-6.4379	177.4358	46.4417
P_{51}	-14.771	-22.928	-6.269	43.144	535.504	-17.3663	-23.941	-9.0606	41.6418	532.688
P_{52}	36.985	-25.978	22.976	99.996	560.691	40.2864	-20.9209	25.6725	104.5692	563.7259
P_{53}	-4.865	2.452	3.670	37.692	154.793	-6.192	3.4425	0.4258	38.2139	153.0525