Detection of Yellow Starthistle through Band Selection and Feature Extraction from Hyperspectral Imagery

Xin Miao, Peng Gong, Sarah Swope, Ruiliang Pu, Raymond Carruthers, and Gerald L. Anderson

Abstract

To effectively display hyperspectral imagery for visualization purposes, the three RGB channels should be selected or extracted from a hyperspectral image under the criteria of maximum information or maximum between-class separability. Seven band selection (OIF, SI, CI, divergence, transformed divergence, B-distance, JM-distance) and five feature extraction (principal component analysis, linear discriminant analysis, class-based PCA, segmented PCT (SPCT), independent component analysis) methods and their variations are examined and compared using CASI *hyperspectral imagery with the goal of detecting* Centaurea solstitialis *(yellow starthistle or YST), an invasive weed, in an annual grassland in California. Three indicators, information index (Infodex), separability index (Sepadex) and average correlation coefficient (ACC) are proposed to evaluate the quality of the generated images. The results suggest that both the combination of the three* SPCT *channels and the combination of the second PCA channel with the positive and negative of the first LDA channels (PCA2, LDA1, -LDA1) can enhance our ability to visualize the distribution of YST in contrast to the surrounding vegetation.*

Introduction

Invasive species pose one of the greatest threats to native biodiversity worldwide (Mooney and Drake, 1987; Vitousek, 1990). They are the leading cause of species endangerment, second only to outright habitat destruction (Wilcove *et al.*, 1998) and often cause irreversible alterations to ecosystems. Therefore, mapping the distribution and abundance of invasive species across large areas is essential to the effective management of our natural resources as well as to improving our understanding of the ecology of biological invasions. Remote sensing plays an important role in invasive species management because it allows us to map invasive species across large spatial extents often with greater accuracy and precision than commonly used field-based methods. It will likely become even more important to the management of invasive species as we improve upon our ability to extract data from existing technologies. Because of its exceptional spectral resolution, hyperspectral imaging offers great promise in mapping the abundance of particular species over large spatial extents. For example, several recent publications have applied hyperspectral image analysis to the task of mapping invasions and even individual invasive species (Ustin *et al.*, 2001; Lass *et al.*, 2002; Underwood *et al.*, 2003), with varying levels of success.

However, hyperspectral imagery tends to be more difficult to process than the commonly used multispectral imagery due to the geometrical and statistical properties of high dimensional data (Landgrebe, 2002). Hyperspectral image visualization techniques are among the effective ways to explore the characteristics of hyperspectral image data at the early stage of image analysis. However, only a few studies on hyperspectral image visualization in the broad sense can be found (Gong *et al.*, 2001; Polder and Heijden, 2001; Hsu, 2002). We will focus on the evaluation of band selection and feature extraction in this paper.

Band selection and feature extraction serve essentially dimension reduction purposes. In this research, we aim at keeping only three bands or feature channels as the output channels for image display and image interpretation. The criteria of maximum information or maximum between-class separability were used in band selection and feature extraction. Three band selection methods, Optimum Index Factor (OIF), Sheffield index (SI), and an alternative index (CI) have been introduced to contain the most information among available bands (Beaudemin and Fung, 2001). Four separability measures: divergence, transformed divergence, Bhattacharyya distance, and Jeffreys-Matusita distance were applied to determine the best band subset of a multispectral

Xin Miao is with the Department of Geography, Geology and Planning, Missouri State University, Springfield, MO 65897 and with the Department of Environmental Science, Policy and Management, University of California at Berkeley (xinmiao@missouristate.edu).

Peng Gong is with the Department of Environmental Science, Policy and Management, University of California at Berkeley, Berkeley, CA 94720 and with the State Key Lab of Remote Sensing Science, Chinese Academy of Sciences and Beijing Normal University, Beijing, China, (gong@nature.berkeley.edu).

Ruiliang Pu is with the Department of Environmental Science, Policy and Management, University of California at Berkeley, Berkeley, CA 94720.

Sarah Swope and Raymond Carruthers are with the Agricultural Research Service, USDA, Albany, CA 94710.

Gerald L. Anderson is with the USDA Agricultural Research Service, Northern Plains Area, Sidney, MT 59270.

Photogrammetric Engineering & Remote Sensing
Vol. 73, No. 9, September 2007, pp. 1005–1015.

0099-1112/07/7309–1005/$3.00/0

video system (Mausel *et al.*, 1990). They are essentially similar. Principal component analysis (PCA) is an important technique in image enhancement and display (Richards, 1993). These conventional approaches have often been used in multispectral image visualization (Bellec and Legleau, 1992; Clevers and Vanstokkom, 1992; Haack and Jampoler, 1995). However, other feature extraction methods, such as linear discriminant analysis (LDA), class-based PCA (CPCA), segmented PCT (SPCT), and independent component analysis (ICA) may also be potential candidates for image visualization (Okada and Tomita, 1985; Cardoso and Souloumiac, 1993; Jia and Richards, 1999; Xu and Gong, 2005; Miao *et al.*, 2007), although they have not been broadly applied in the visualization of hyperspectral imagery.

Conventional band selection and feature extraction methods largely fall into two categories in terms of criteria and objectives: OIF, SI, CI, and PCA attempt to keep maximum image variability (regarded as information), and they are primarily used in preliminary exploration and image compression; the four class separability measures and LDA attempt to maximize the between-class separability, and they are primarily used for image classification (Mausel *et al.*, 1990; Richards, 1993; Webb, 2002). Additionally, SPCT, ICA, and CPCA have unique assumptions and characteristics (Cardoso and Souloumiac, 1993; Jia and Richards, 1999; Miao *et al.*, 2007). How to apply these methods appropriately in hyperspectral image visualization is a challenge. Regardless of the method(s) chosen, the principle is to flexibly tailor and synthesize these algorithms to adapt to a particular application.

Yellow starthistle (*Centaurea solstitialis* or YST) is one of the most problematic invasive species in California (Lass *et al.*, 1996; Lass *et al.*, 2000). It is now estimated to infest 15 to 20 million acres in California. It can access to deep soil moisture during the dry summer. Because of its spiny yellow-flowered heads, livestock and wildlife avoid grazing in heavily infested areas. It is also highly toxic to horses (DiTomaso and Gerlach, 2000). In this study, we applied band selection and feature extraction methods to CASI hyperspectral imagery for visualization with the purpose of detecting this invasive weed at a study site in the Central Valley grasslands of California. The remainder of this paper is organized as follows. The methodology of band selection and feature extraction is described in the next section. Then, the performance of the band selection and feature extraction methods for CASI image visualization are compared and evaluated. Finally, the experimental results and discussion of the results are presented.

Background

Band Selection

Maximum Information Criteria

Optimum Index Factor (OIF), Sheffield index (SI), and an alternative index (CI) are used to rank band subsets according to their information content (Beaudemin and Fung, 2001):

$$OIF = \sum_{i=1}^{3} SD_i \Big/ \sum_{i=1}^{2} \sum_{j=i+1}^{3} |r_{ij}| \tag{1}$$

$$SI = |M_{3\times3}| \tag{2}$$

$$CI = |M_{3\times3}| \Big/ \prod_{i=1}^{3} m_{ii} = |R_{3\times3}| \tag{3}$$

where SD_i is the standard deviation of band *i*, M is the covariance matrix, R is the correlation matrix, m_{ij} and r_{ij} are the elements of M and R, respectively. OIF and SI tend to select those bands with high variances and low pair-wise correlation. Beaudemin *et al.* (2001) illustrated that OIF and SI were different only from weights of individual band variances. Then, a more robust index, CI was proposed, which measured how the data subset fills the space defined by the size of its projection on each axis in the original bands. Actually, CI is equivalent to the determinant of the correlation matrix. Therefore, it can be understood as the normalized version of SI.

Maximum Between-class Separability

Divergence D_{ij} is a commonly used separability measure of band selection for classification (Mausel *et al.*, 1990; Richards, 1993). For the multi-normal distributions the divergence between two classes is:

$$D_{ij} = \tfrac{1}{2} Tr[(\Sigma_i - \Sigma_j)(\Sigma_j^{-1} - \Sigma_i^{-1})] + \tfrac{1}{2} Tr[(\Sigma_i^{-1} + \Sigma_j^{-1})(m_i - m_j)(m_i - m_j)^T] \tag{4}$$

where Σ_i is the class covariance matrix, m_i is the class mean vector. Since D_{ij} increases quadratically with increasing class separation, which is misleading in terms of classification, Transformed Divergence (TD) has been proposed to enforce a saturating behavior (Richards, 1993):

$$TD_{ij} = 2(1 - e^{-d_{ij}/8}) \tag{5}$$

where TD_{ij} ranges from 0 to 2, and the saturation value 2 indicates the maximum spectral separability.

Bhattacharyya distance (B-distance) is another measure of class separability (Kailath 1967):

$$B_{ij} = \tfrac{1}{8}(m_i - m_j)^T \left[\frac{\Sigma_i + \Sigma_j}{2}\right]^{-1} (m_i - m_j) + \tfrac{1}{2} \ln \left[\frac{|(\Sigma_i + \Sigma_j)/2|}{|\Sigma_i|^{1/2}|\Sigma_j|^{1/2}}\right]. \tag{6}$$

A similar transformed form of B-distance is referred to as Jeffreys-Matusita distance (JM-distance) (Richards, 1993):

$$J_{ij} = 2(1 - e^{-B_{ij}}). \tag{7}$$

The Equations 4 through 7 are pair-wise measures of separability. For a given measure, the average value of all possible class pairs can be used to represent the overall performance of separability. TD is a common separability measure due to its computational efficiency comparing with JM-distance. However, JM-distance was reported to yield slightly better results for optimal band selection than TD (Mausel *et al.*, 1990).

Feature Extraction

Maximum Information Criteria

Principal component analysis (PCA) provides a sequence of best linear approximations for a set of high dimensional data and data can be represented without correlation (Richards, 1993). Principal axes are the eigenvectors of the sample covariance matrix. Let eigenvalues be indexed in the order of decreasing magnitude, and θ correspond the first p eigenvectors. The projection process can be described as:

$$\hat{X} = \theta^T (y - m) \tag{8}$$

where y is the q-dimension data, $\hat{X}$ is the estimated p-dimension data, $p < q$, and m is the mean vector.

Maximum Between-class Separability

Linear Discrimant Analysis (LDA) is a feature reduction routine especially for high-dimensional data classification. Conventionally, LDA can find a set of new axes so that the ratio of

between-class variance to within-class variance is as large as possible. The classical Fisher linear discriminant is a generalized eigenvalue problem (Richards, 1993; Yu *et al.*, 1999):

$$(\Sigma_B - \lambda\Sigma_W)d = 0 \tag{9}$$

where Σ_B and Σ_W are between-class and within-class covariance matrix, respectively. However, if n classes are defined, only $n - 1$ eigenvectors exist and they are not orthogonal. It is reported that an improved orthogonal system is more powerful than the classical discriminant vectors (Okada and Tomita, 1985; Duchene and Leclercq, 1988). The problem is defined as:

$$\text{Max: } \frac{d_n'\Sigma_B d_n}{d_n'\Sigma_W d_n} \tag{10}$$

but with the constraints: $d_1'd_n = d_2'd_n = \cdots = d_{n-1}'d_n = 0$. The first vector of orthogonal LDA vector is also the first Fisher linear discriminant vector. Furthermore, they proposed a generalized PCA by defining a *best variance* vector d_n, orthogonal to a set of previously computed orthogonal discriminant vectors $d_1 \cdots d_{n-1}$. The *best variance* vector d_n is the eigenvector of:

$$N = (I - d_1d_1' - \cdots - d_{n-1}d_{n-1}')T \tag{11}$$

where T is the total covariance matrix. The orthogonal transformation is a rotation without modifying the initial shape of the data.

Other Feature Extraction Methods

There are three other feature extraction methods with different assumptions and characteristics. Segmented Principal Component Transform (SPCT) can be regarded as a compromise between band selection and global PCA. SPCT employs the block structure of the hyperspectral image correlation matrix to group the adjacent spectral bands, and then apply PCA only in each block. This method not only reduces the dimension of the data but also keeps the trail of band information and increases interpretability (Jia and Richards, 1999). Class-based PCA (CPCA) mainly applies feature extraction on each individual land-cover class (Miao *et al.*, in review). The philosophy of CPCA is to focus information extraction only on some specific classes that we are interested in. Independent Component Analysis (ICA) is a blind signal separation method. The model of ICA can be described as:

$$x = As \tag{12}$$

where $A \in R^{n\times n}$ is an unknown non-singular mixing matrix, and n independent non-Gaussian signals are collected in $n \times 1$ vector s. Without knowing the source signals and mixing matrix, we need to recover the original signals from the observation x (Cardoso, 1994). ICA has already been used in target detecting and image classification for hyperspectral imagery (Chang *et al.*, 2002).

Information Index and Separability Index

In order to evaluate the quality of the generated image under a unified framework, we propose three indicators: information index (Infodex), separability index (Sepadex), and average correlation coefficient (ACC).

$$Infodex = \sum_{j=1,2,3} \lambda_j' \Big/ \sum_{i=1,\ldots n} \lambda_i \tag{13}$$

$$Sepadex = \frac{1}{3}\cdot \sum_{i=1,2,3} \Sigma_{Bi}'/\Sigma_{Wi}' \tag{14}$$

$$ACC = \frac{1}{3}\cdot \sum_{i=1,2,3} |r_i| \tag{15}$$

where λ_i is the eigenvalue of the covariance matrix of the hyperspectral image data, λ_j' is the eigenvalue of the covariance matrix of three output channels (before linear stretching enhancement), n is the total number of hyperspectral image channels, Σ_{Bi}' and Σ_{Wi}' are between-class and within-class covariance matrix of three output channels, and r_i is the pair-wise correlation coefficient among output channels. Clearly, Infodex represents the ratio of the variance (information) of the generated image to the total variance of the original hyperspectral image; Sepadex represents the average ratio of between-class variance to within-class variance of the output channels; and ACC represents the correlations among three output channels. These three indicators reflect the information capacity, class separability and information redundancy of the generated image. Ideally, a good image should contain more information, show higher class separability and have lower correlations among three output channels. Infodex also has some relation with SI: SI actually is the product of the three eigenvalues of the covariance matrix of output channels $\prod_{i=1,2,3} \lambda_i'$, while the numerator of Infodex is the sum of them $\sum_{i=1,2,3} \lambda_i'$.

Methods

Study Site

Our study site was located at the western edge of California's Central Valley grasslands (39°00′ N, 122° 21′ W), approximately 130 km northeast of San Francisco. The site is dominated by annual grasses soft brome (*Bromus hordeaceus*), wild oats (*Avena fatua*), and medusahead (*Taeniatherum caput-medusa*). Yellow starthistle (*Centaurea solstitialis*) has been present at the site for decades. The study area is largely a flat valley bottom with very little topographic variation except at the far eastern edge where it transits to oak woodland on steep (>30 percent) slopes. The soil is a loam and elevation ranges from 314 m to 380 m. The land is privately owned and has not been grazed by domestic livestock or treated with prescribed fire or herbicides since prior to 1998. The patches of live yellow starthistle vary in distribution from year to year. The dynamics are probably caused by annual variation of climate and site lighting conditions in the study area.

Data Acquisition and Preparation

CASI-2 (Compact Airborne Spectrographic Imager-2, referred to as CASI in this paper) is a 12-bit charge coupled device (CCD) pushbroom imager designed for the acquisition of visible and near infrared hyperspectral imagery. Two CASI hyperspectral images were used to examine and verify the effectiveness of image visualization algorithms. One image was captured on 15 July 2003. It has 36 channels; however, channels 1 to 3 and 30 to 36 are obviously blurred due to the decreasing of spectral response at the two ends of the spectral range (Milton and Choi, 2004). Therefore, we only selected channels 4 to 29 covering a wavelength range from 465 nm to 847 nm in our study. The spatial resolution of this image is 3 m. After geometric correction, a 700 pixel × 350 pixel image was cropped from the whole scene as the study area shown in Figure 1. The other CASI image was

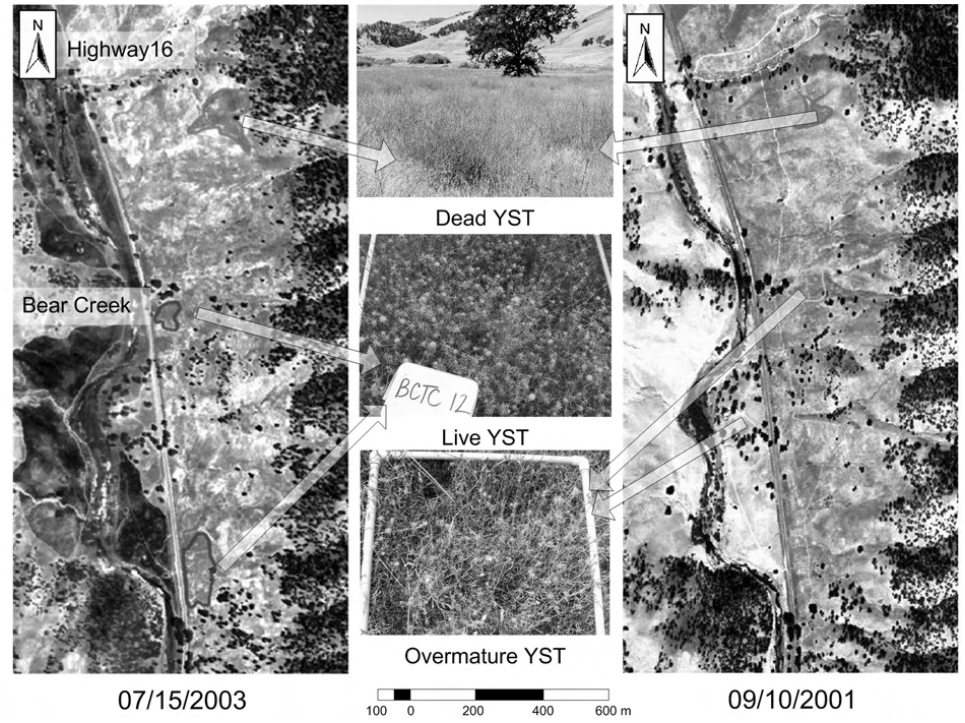

Figure 1. CASI images cropped from the Bear Creek scene, California (RGB versus 646 nm, 525 nm, and 480 nm) and YST patterns.

taken in the same area on 10 September 2001. It has 48 channels, but we only kept channels 3 to 42 with a wavelength range from 450 nm to 896 nm due to the blurring problem; the spatial resolution is 2 m. There is a common problem for both CASI images: a gradual horizontal shift at the total amplitude of one pixel exists across from first channel to the last channel. This inter-channel registration problem could produce some blurring to the generated images. We used the image taken on 15 July 2003 to evaluate the band selection and feature extraction algorithms for image visualization and the image taken on 10 September 2001 to further test the robustness and effectiveness of our algorithms.

Our first step was to generate three channels to enhance the visualization of areas infested with YST in the image taken on 15 July 2003. Six land-cover classes were identified in this CASI image: live YST (at full flower stage), dead YST (senesced plant materials from last year), riparian vegetation (willows, cottonwoods, and saltcedar), native oak trees, background material (soil and short dead annual grasses), and an area subjected to prescribed burning by the Bureau of Land Management to the west of our study site about two weeks prior to image capture. The numbers of all class samples are listed in Table 1 and class spectral signatures are plotted in Figure 2. These class samples were used to estimate the separability measures (divergence, transformed divergence, B-distance, and JM-distance) and LDA.

The original CASI images have a 12-bit depth. We transformed all images to 8-bit depth for computing and display convenience. It should not seriously affect the visual effect due to limited color-resolving power of the human eye (Richards, 1993). To highlight the visual contrast in each selected channel, a linear stretching enhancement technique has been used to

TABLE 1. NUMBER OF CLASS SAMPLES FOR THE CASI IMAGE TAKEN ON 15 JULY 2003

Class	1. Live YST	2. Dead YST	3. Riparian Vegetation	4. Oak	5. Background Material	6. Burn Area
Sample	678	926	423	419	622	2450

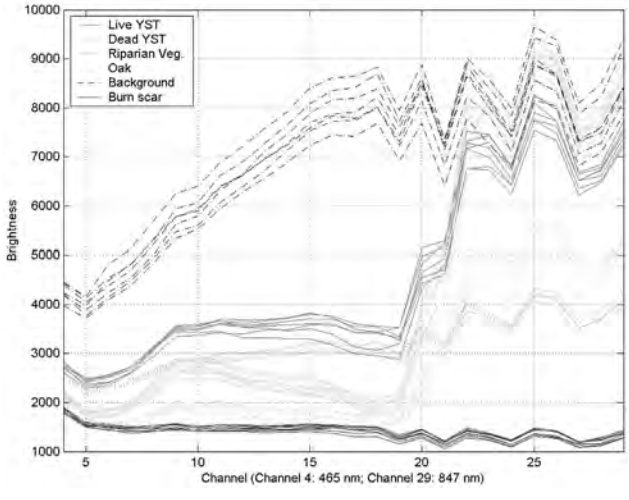

Figure 2. Spectral signatures of classes.

show all images in this paper (Beaudemin and Fung, 2001). Another problem is how to assign the RGB colors to the three selected bands (channels). Haack *et al.* (1995) suggested to try all six combinations and to select the best image subjectively. Here we adopt another method: if the three channels are derived from band selection or SPCT, we assign B, G, and R to the bands in the increasing order of wavelength, so that the false color image can abide by the conventions of photo interpretation (Richards, 1993); if the three channels are derived from feature extraction without obvious interpretability, we simply assign R, G, and B to the three feature channels ranking in the order of significance. This has proven to be an effective color assignment in our study.

Finally, two experienced field researchers were invited to evaluate all selected or extracted images subjectively coupling with two-year field data. Our conclusions are partially based on their comments and the results of field survey of YST distribution in our study area from year 2001 to 2003. In order to minimize the subjectivity, we use the information index (Infodex), separability index (Sepadex), and average correlation coefficient (ACC) to rank all generated images. We believe the best visual effect should be a good trade-off among these three indicators. In addition, we prepared two land-cover classification maps corresponding to the two CASI images. The two maps were produced by evenly selecting seven bands from all available bands and applying the supervised maximum likelihood classification algorithm (Richards, 1993). A half of the class samples in Table 1 were randomly selected to serve as training samples, and the other half of the class samples were used to estimate the accuracies of classification. The overall accuracies of classification are 85 percent for the image taken on 15 July 2003 and 89 percent for the image taken on 10 September 2001. These two maps may serve as a general reference to judge the visual effect of the YST patches in the generated image.

Band Selection

Three band selection indices, OIF, SI, and CI and four separability measures, divergence, transformed divergence, B-distance, and JM-distance have been applied to select three bands from the CASI image taken on 15 July 2003[1]. Since our major objective was to detect YST, two sets of average between-class separability measures have been compared. First, we averaged all pair-wise measures (a total of 15 measures); then we only averaged those pair-wise measures related to live and dead YST (a total of nine measures). The ten best band subsets for each method are listed in Table 2. Since the three band selection indices and four separability measures are based on different criteria, two histograms are plotted to count the frequencies of those best bands listed in Table 2, as shown in Figure 3.

Figure 3a shows that bands 4, 18, 25, and 29 have higher frequencies. This fact is also indicated by the best OIF, SI, and CI band subsets in Table 2. Comparing four possible three-band subsets among bands 4, 18, 25, and 29, we find bands 4, 18, and 29 provide the best visual effect of YST in contrast to its surroundings. From Figure 3b, bands 9, 18, 23, 27, and 29 have higher frequencies. Comparing ten possible three-band combinations among them, bands 9, 18, and 27 provide the best visual effect of YST. According to the comments of field researchers and comparing with the field survey data, the latter shows a clearer live and dead YST patches in contrast to surroundings (Plate 1a).

[1]All results in this paper were acquired by programming in Matlab 6.5.

Feature Extraction

PCA and LDA are commonly used feature extraction methods. The first three PCA channels can explain nearly 99 percent of the total data variance of the CASI image taken on 15 July 2003. However, the edges of yellow starthistle appear blurred in the image generated by the first three PCA channels. We expect that LDA can provide a clearer visualized image. However, the image composite of the first three classical Fisher LDA channels has no obvious improvement. The orthogonal LDA is also problematic: the first three orthogonal LDA channels have strong correlations (correlation coefficients are almost 1) among them and produce a nearly grey-level composite image. Then, we tried to generate a *best variance* vector perpendicular to the first two orthogonal LDA vectors. But, the image generated by the first two orthogonal LDA channels and a *best variance* channel is also an almost grey-level image.

We summarize the angles between pair-wise vectors and correlation coefficients between pair-wise channels in Table 3. First, we found that the correlation coefficient between two channels have no direct relation with the angle between two vectors. For example, the channels corresponding to the first and second orthogonal LDA vectors have a strong correlation coefficient (close to 1). This is due to the geometric characteristics of hyperspectral space: two orthogonal vectors may share a similar pattern, so the local difference between the vectors may be absorbed by the global similarity of their trends. Second, the first LDA channel and the first PCA channel have a strong correlation coefficient (0.99). It seems to suggest that classes can also be largely separated by projecting them onto the first PCA vector. This also verified the claim that classes are frequently distributed in the direction of maximum data scatter in remote sensing (Richards, 1993). Third, the *best variance* vector is very close to the first PCA vector with the angle of 8.5°; the reason is that the *best variance* vector is the first eigenvector of

$$N = (I - d_1 d_1' - d_2 d_2')T \tag{16}$$

where d_1, d_2 are 26-dimensional normalized orthogonal LDA vectors, $d_1 d_1'$ and $d_2 d_2'$ are negligible, therefore, $N \simeq T$, and the *best variance* vector is very close to the first PCA vector.

Although the image generated from the first three channels of PCA, classical Fisher LDA, and orthogonal LDA seems discouraging, we noticed that if we force both the Red and the Green channels to be the first LDA channel, and the Blue channel to be its *negative* channel by projecting data on the reverse of the first LDA vector (RGB versus LDA1, LDA1, -LDA1), we generated almost a two-color image as shown in Plate 1b. We can identify clearly from Plate 1b that the pale yellow color patches on the east side of the creek are primarily dead YST. This phenomenon has three reasons. First, the yellow color comes from the equal mixture of the Red and Green channels, which is the complement color of blue. Second, the classes have the highest separability when projecting on the first LDA vector. Third, a class with a high digital value in LDA1 channel will have a low digital value in the negative LDA1 channel, so the contrast is enhanced. Therefore, the pale yellow dead YST patches are clearly emerged on the blue background. However, this image seems to lose much variance of data due to the strong correlation among three channels. Therefore, we attempted to combine the advantages of PCA by selecting the second principal channel for Red channel as shown in Plate 1c. We did not select the first PCA channel because it has a strong correlation with LDA1 channel (Table 3). Although live YST (pale purple color patches on the right side of the highway) and riparian vegetation along the creek are still difficult to separate from

TABLE 2. BEST TEN BAND COMBINATIONS FOR THE CAST IMAGE TAKEN ON 15 JULY 2003

OIF[1]	SI	CI	D	TD	B-distance	JM-distance
18,25,26	18,23,28	4,18,25	19,20,23 (19,21,23)[2]	11,27,29 (8,18,27)	19,20,23 (19,22,23)	14,27,29 (9,17,28)
17,18,25	18,22,28	4,18,26	18,20,23 (18,21,23)	16,18,29 (8,18,26)	18,20,23 (18,20,23)	15,27,29 (9,17,27)
18,25,29	17,23,28	4,18,27	17,20,23 (19,20,23)	9,11,29 (8,18,25)	19,23,27 (19,20,23)	6,10,23 (9,19,29)
18,25,29	18,23,29	4,18,23	19,21,23 (18,20,23)	12,28,29 (8,18,29)	20,23,24 (18,22,23)	6,10,25 (9,18,29)
16,18,25	16,23,28	4,18,24	16,20,23 (17,21,23)	9,27,29 (8,18,28)	20,22,23 (18,23,27)	16,27,28 (9,19,27)
17,25,26	17,22,28	4,18,29	19,23,27 (17,20,23)	11,27,28 (7,23,24)	16,20,23 (19,23,28)	13,27,28 (9,18,27)
18,25,29	18,22,29	4,18,28	18,21,23 (16,21,23)	6,10,29 (8,18,23)	17,20,23 (19,23,27)	14,27,28 (11,17,27)
18,23,25	17,23,29	4,19,25	19,23,28 (19,22,23)	12,27,28 (8,17,27)	19,23,27 (19,23,29)	6,11,29 (9,16,27)
17,18,26	18,25,29	4,19,26	15,20,23 (16,20,23)	6,11,29 (8,18,29)	19,23,29 (18,23,29)	6,10,24 (9,15,27)
18,26,29	16,22,29	4,19,27	14,20,23 (18,22,23)	16,18,27 (8,17,26)	19,23,27 (19,21,23)	6,11,27 (9,15,29)

Note 1. OIF (Optimum Index Factor), SI (Sheffield Index), and CI (an alternative index) are three indices based on the maximum information criteria; D (divergence), TD (transformed divergence), B-distance (Bhattacharyya distance) and JM-distance (Jeffreys-Matusita distance) are four measures based on the maximum between-class separability.
Note 2. Band combinations in the parentheses are derived from averaged pair-wise measures only related to YST.

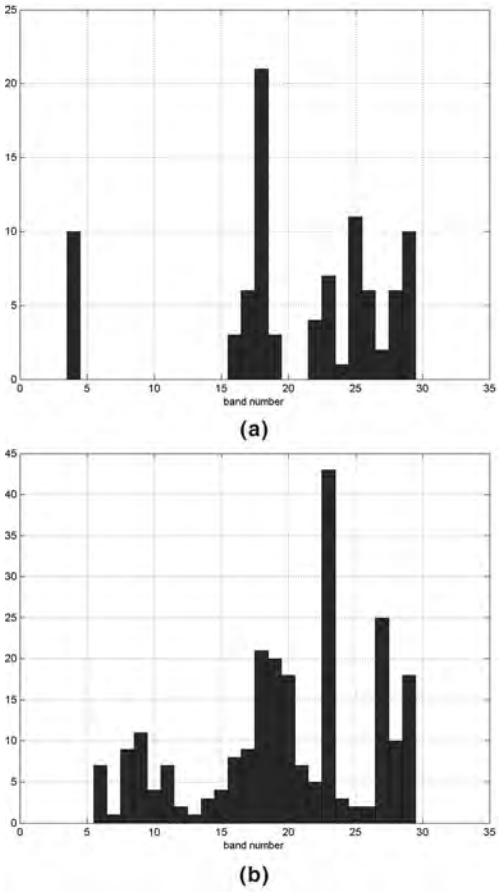

Figure 3. Histograms of best bands from Table 2: (a) OIF, SI, and CI based on the maximum information criterion, (b) divergence, transformed divergence, B-distance, and JM-distance based on the maximum between-class separability criterion.

color difference, we can be reasonably certain which species are which based on their habitat preferences: the riparian species will, by definition, be confined to the riparian corridor, and YST is an arid grassland species. The dead YST patches (pale green) on the right side of the highway clearly emerge from the blue background. Furthermore, the oak trees that dominate the eastern edge of the image (yellow dots) are easily distinguished from background soil, annual grasses, and dead YST. This image not only exhibits the complex land-cover variance but also clearly illustrates the areas of live and dead YST. Like the PCA image, this image is also a little blurred. Since the three orthogonal LDA channels and the first PCA channel have strong correlations, we may also use them instead of LDA1 in this case, and the visual effects are similar. However, we found that LDA1 provides a slightly better visual effect for dead YST in contrast to the surroundings.

One of the shortcomings of the PCA or LDA is that the interpretability is sacrificed because the output channels are the mixture information of spectral bands. SPCT can compensate for the loss of interpretability (Richards, 1993). The block structure of the correlation matrix of the CASI image taken on 15 July 2003 indicates a border at band 20 with a wavelength of 708 nm, which is close to the vegetation *red edge* location (Pu *et al.*, 2003). We then grouped bands 4 to 12 (465 to 585 nm), bands 13 to 19 (601 to 692 nm), and bands 20 to 29 (708 to 847 nm), respectively. The first group represents the blue-green bands; the second one represents red bands and the third one represents near infrared bands. Then, PCA was applied to each group. The first principal channel of each of the three groups is used to compose an image as shown in Plate 1d, which is similar to Plate 1a but exhibiting more color variance.

If PCA is only applied to live and dead YST classes, the principal axes can represent the directions of maximum variance of YST classes, and the data variance from other classes can be potentially reduced. This is the principle of Class-based PCA (CPCA). The live and dead YST class samples are taken from Table 1. Although CPCA may suffer from the poor representation of class samples when dealing with hyperspectral imagery, it is reasonable in our case since the live and dead YST samples are close to or more than 20 times of the number of spectral bands (a total of 26 bands). The relation between the first CPCA vector and other vectors are illustrated in Table 3. However, the images composed from the first three principal channels of CPCA for both dead and live YST does not show clear patches. Most probably the reason is the CPCA vectors do not suppress the variance of other classes in this case.

ICA is another high-dimensional data analysis technique. We adopt the algorithms developed by Cardoso and Souloumiac (1993). Without prior knowledge, we simply assume four independent components existing in the CASI image and generate four output channels. Although it is

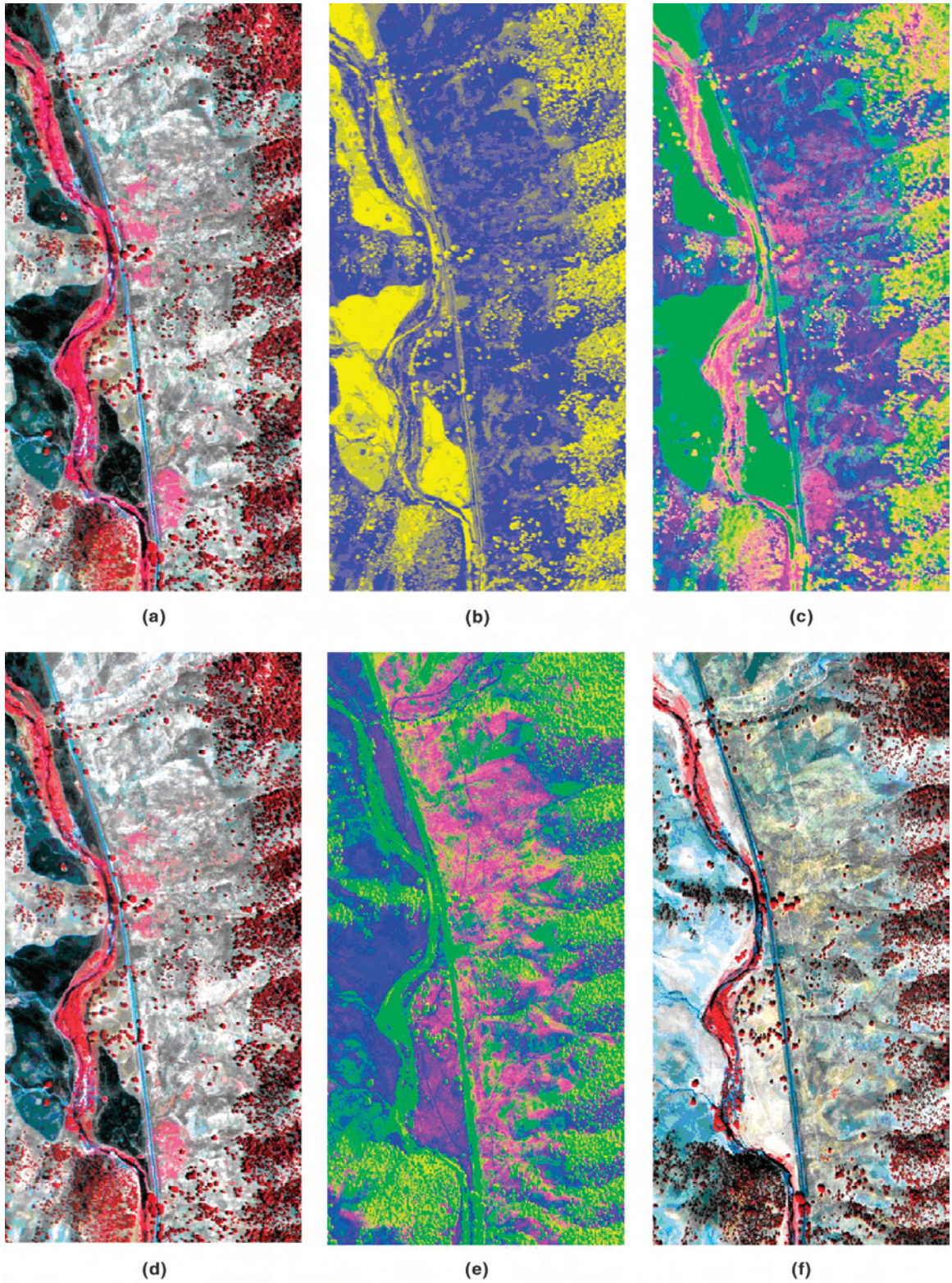

Plate 1. False color images: (a) derived from band selection algorithms of divergence, transformed divergence, B-distance, and JM-distance for the CASI image taken on 15 July 2003 (Figure 3b), RGB versus band 27 (816 nm), band 18 (677 nm), band 9 (540 nm), (b) RGB versus LDA1 and-LDA1 for the CASI image taken on 15 July 2003, (c) RGB versus PCA2, LDA1, and-LDA1 for the CASI image taken on 15 July 2003, (d) RBG versus the first principal channels of the three SPCT segments for the CASI image taken on 15 July 2003, (e) RGB versus PCA2, LDA1, -LDA1 for the CASI image taken on 10 September 2001, and (f) RGB versus the first principal channels of the three SPCT segments for the CASI image taken on 10 September 2001.

TABLE 3. ANGLES AND CORRELATION COEFFICIENTS BETWEEN PAIR-WISE VECTORS BASED ON FEATURE EXTRACTION FOR THE CASI IMAGE TAKEN ON 15 JULY 2003

	PCA1	PCA2	PCA3	Orth LDA1	Orth LDA2	Orth LDA3	LDA2	LDA3	LDA–BV	CPCA1	ICA4
PCA1	–	0	0	0.99	0.99	0.99	0.49	0.21	1.00	0.93	0.95
PCA2	90	–	0	0.01	0.03	0.01	0.85	0.30	0.00	0.37	0.06
PCA3	90	90	–	0.10	0.09	0.10	0.11	0.30	0.00	0.01	0.26
OrthLDA1	85.4	89.9	85.7	–	1.00	1.00	0.48	0.16	0.99	0.94	0.98
OrthLDA2	82.6	89.5	83.2	90	–	1.00	0.46	0.17	0.99	0.93	0.98
OrthLDA3	78.4	89.7	78.9	90	90	–	0.48	0.17	0.99	0.93	0.98
LDA2	88.1	82.9	86.0	80.5	68.9	84.1	–	0.23	0.49	0.14	0.44
LDA3	89.6	89.2	86.7	72.9	87.1	85.0	67.9	–	0.21	0.08	0.07
LDA-BV	8.5	89.9	88.8	90	90	78.4	86.3	88.8	–	0.93	0.95
CPCA1	42.5	48.0	86.5	86.9	84.7	81.8	86.4	89.9	42.9	–	0.90
ICA4	74.2	86.8	37.1	84.2	80.7	75.2	87.6	86.2	74.7	78.7	–

Note 1. Numbers in the lower left corner of the table indicate the angles (°) between pair-wise vectors; numbers in the upper right corner indicate the absolute value of correlation coefficients.
Note 2. OrthLDA 1 through 3 are orthogonal LDA channels (OrthLDA 1 is also the same as the first Fisher LDA channel); LDA 2, LDA 3 are Fisher LDA channels; LDA-BV is the *best variance* channel following OrthLDA 1 to 2; CPCA 1 is the first class-based PCA channel; ICA 4 is the fourth ICA channels.

difficult to interpret these independent components, the first channel in this case seems to be the high-frequency components, the second channel adds some details on the first channel, the third channel mainly emphasizes on the green vegetation components without showing the dead YST patches, and the fourth channel may be understood as the overall data variance. However, combining any three of the four channels generates an image composite with lots of noise. Another phenomenon is that the fourth independent component is similar to the first PCA channel with a large correlation coefficient of 0.95 (Table 3). Therefore, if we combine the fourth independent component channel and the positive and negative of the first LDA channels, the image composite is similar to Plate 1c. But, the overall visual effect is not better than Plate 1c.

Evaluation of Visual Effects

In order to minimize the subjectivity, we calculated the information index (Infodex), separability index (Sepadex) and average correlation coefficient (ACC) for all image composites (Table 4). It can be noticed that (a) An image with a high Infodex usually has a low Sepadex, or vice versa; (b) The images based on band selection have a lower Infodex and Sepadex comparing with the images derived from feature extraction; (c) The first three orthogonal LDA channels provide the highest Sepadex (72.1), but the Infodex is low (0.05) and the ACC is high (0.99); (d) The classical Fisher LDA has an even lower Infodex (0.01), lower Sepadex and much lower ACC than the orthogonal LDA; and (e) The third SPCT channel that represents the near-infrared band has the highest Sepadex among the three SPCT output channels. Since our objective is to detect live and dead YST patches, the criterion of maximum class separability is a necessary condition. Therefore, PCA and CPCA are not good choices, although they have very high Infodex and low ACC. However, the image combining two orthogonal LDA channels and a *best variance* channel is nearly a grey-level image due to the high correlations among them (ACC is 0.99), although they have high Infodex and Sepadex. This suggests that we need to consider all three indicators when judging the visual effect. We found that the image derived from SPCT and the image combing the second PCA channel and the positive and negative of the first LDA channel (PCA2, LDA1, -LDA1) provide a good trade-off among information capacity, between-class separability and information redundancy (Plate 1c and 1d). The SPCT image emphasizes more on the information capacity, and the image

TABLE 4. INFODEX (INFORMATION INDEX), SEPADEX (SEPARABILITY INDEX) AND ACC (AVERAGE CORRELATION COEFFICIENT) OF ALL GENERATED IMAGES (15 JULY 2003)

		Sepadex (rank)				
RGB Channels	Infodex (rank)	Channel 1	Channel 2	Channel 3	Average (rank)	ACC (rank)
PCA 2,PCA 1,-PCA 1	1.81 (1)	34	63.6	63.6	53.7 (4)	0.33 (5)
PCA 1-3	0.99 (2)	63.6	34	0.82	32.8 (10)	0 (1)
CPCA 1-3	0.99 (3)	58.6	48.1	2.7	36.5 (8)	0.30 (3)
SPCT 1-3	0.97 (4)	41.1	47	59.8	49.3 (5)	0.79 (8)
OrthLDA 1, 2 & Best Variance	0.83 (5)	75.6	71.1	63.3	70.0 (2)	0.99 (9)
PCA2, LDA1, -LDA1	0.20 (6)	34	75.6	75.6	61.7 (3)	0.33 (4)
Bands (1,18,31)[1]	0.10 (7)	2.9	44.1	58.8	35.3 (9)	0.54 (6)
Bands (9,18,27)[2]	0.10 (8)	45.2	44.1	56.3	48.5 (6)	0.75 (7)
OrthLDA 1-3	0.05 (9)	75.6	71.1	69.6	72.1 (1)	0.99 (10)
Fisher LDA 1-3	0.01 (10)	75.6	33.9	2.7	37.4 (7)	0.29 (2)

Note 1. Derived from OIF, SI, and CI (maximum information criteria).
Note 2. Derived from divergence, transformed divergence, B-distance, and JM-distance (maximum between-class separability criteria).

combining PCA2 and LDA1 channels emphasizes more on the between-class separability (Table 4). On the other hand, the image combing the second PCA channel and the positive and negative of the first PCA channel (PCA2, PCA1, -PCA1) is similar to Plate 1c, but the visual separability is not as high as Plate 1d. This image has an Infodex larger than 1 (1.81) because of the coexistence of both the positive and negative of the first PCA channels.

Test the Robustness of Algorithms

According to our research, we propose to combine one channel from PCA with the positive and negative of the first LDA channels (PCA 2, LDA1, -LDA1) to enhance the visual effect of distribution of live or dead YST. SPCT is also an effective visualization approach to detect YST. They both have advantages over the band selection methods in terms of contrasting YST to their surroundings. To test the robustness and effectiveness of these two algorithms, the CASI image taken on 10 September 2001 (Figure 1) was used. It has a finer spatial resolution (2 m) and more spectral bands (40) than the image taken on 15 July 2003. Because this image was gathered when the YST was beginning to senesce, live YST is now classified as over-mature YST (the plants are alive and producing flowers and seeds, but are beginning to senesce). We need to identify six classes, namely over-mature YST, dead YST, riparian vegetation, native oak trees, background material (soil and short dead annual grasses), and tall dead grasses mainly distributing on the west of the creek.

The second principal component channel and the positive and negative of the first LDA channel generate a clear YST distribution pattern as shown in Plate 1e. Although some over-mature YST (pale purple) is mixed with the dead YST (pale green) on the east of the highway, they are still distinguishable from the blue background. Furthermore, the low density of the over-mature YST on the west side of the highway is enhanced, which is important for weed control and landscape management. However, like Plate 1c, the image is also a little blurred. Next, the 40 channels were grouped into three subsets according to the structure of correlation matrix: Bands 3 to 17, 18 to 25, and 26 to 42, so that the SPCT image was generated and shown in Plate 1f. Over-mature YST (yellow) and dead YST (dark grey) are shown clearly in Plate 1f. As compared with the classification map and field survey data, Plate 1e and 1f both illustrate the distributions of over-mature and dead YST.

We also produced other false color images for this CASI image based on band selection and feature extraction. The information index, separability index, and average correlation coefficient for these images are summarized in Table 5. The rankings of three indicators in Table 5 are similar to Table 4. The image derived from SPCT and the image combing the second PCA channel and the positive and negative of the first LDA channels both provide good trade-offs among the three indicators.

Discussion and Conclusion

This paper examines and compares the band selection and feature extraction algorithms for hyperspectral image visualization for detecting the invasive weed YST. As for band selection, the four separability measures commonly used for classification provide a slightly better visual effect than OIF, SI, and CI in term of distinguishing YST from its surroundings. The advantage of band selection methods lies in its conservation of physical interpretability. However, band selection is also a time-consuming process. The best band subset has to be selected among $C_{26}^{3} = 2600$ possible band combinations in our first case. The subset number increases dramatically with the increasing number of hyperspectral image bands.

Feature extraction methods can extract useful information under the criteria of maximum information or maximum between-class separability. However, the first three PCA channels blurred the distribution of YST since it only considers the overall data variance. The image derived from classical Fisher LDA has no obvious improvement. The image composed from the first three orthogonal LDA channels is nearly a grey-level image due to the strong correlation among output channels. We propose that combining one PCA channel (the second PCA channel in this case) with the positive and negative of the first LDA channels can combine their advantages and offer a better visual effect. We did not choose the first PCA channel because that channel usually has strong correlation with the first LDA channel, since classes are frequently distributed in the direction of maximum data scatter in remote sensing data. PCA and LDA are relatively easy to calculate, but they generally lose the interpretability due to the mixture of band information. Furthermore, it is difficult to interpret the color assignments corresponding to land-cover classes. Another common problem of feature extraction is that the images are blurred to some extent. This blurring phenomenon might be mainly

TABLE 5. INFODEX (INFORMATION INDEX), SEPADEX (SEPARABILITY INDEX), AND ACC (AVERAGE CORRELATION COEFFICIENT) OF ALL GENERATED IMAGES (10 SEPTEMBER 2001)

		Sepadex (rank)				
RGB channels	Infodex (rank)	Channel 1	Channel 2	Channel 3	Average (rank)	ACC (rank)
PCA 2,PCA 1,-PCA 1	1.68 (1)	35.3	60.2	60.2	51.9 (4)	0.33 (4)
PCA 1-3	0.99 (2)	60.2	35.3	1.97	32.5 (9)	0 (1)
CPCA 1-3	0.99 (3)	55.2	40.1	5.2	33.5 (8)	0.32 (3)
SPCT 1-3	0.96 (4)	40.1	43.2	55.8	46.4 (5)	0.75 (7)
Orthagonal LDA 1, 2 & Best Variance	0.74 (5)	70.8	67.3	60.2	66.1 (2)	0.94 (9)
PCA2, LDA1, -LDA1	0.17 (6)	35.3	70.8	70.8	59.0 (3)	0.45 (5)
Bands (5,21,35)[1]	0.10 (7)	5.2	38.2	51.6	31.7 (10)	0.72 (6)
Bands (10,22,38)[2]	0.11 (8)	32.2	39.1	52.4	41.2 (6)	0.76 (8)
OrthLDA 1-3	0.08 (9)	70.8	67.3	64.2	67.4 (1)	0.99 (10)
Fisher LDA 1-3	0.07 (10)	70.8	34.8	10.6	38.7 (7)	0.28 (2)

Note 1. Derived from OIF, SI, and CI (maximum information criteria).
Note 2. Derived from divergence, transformed divergence, B-distance, and JM-distance (maximum between-class separability criteria).

caused by the re-projection of minor variability in minor component images (Gong, 1993; Pu and Gong, 2004). Besides, the blurring among channels at the short wavelength and near-infrared range due to the decrease of detector sensitivity and the inter-channel registration shift could also cause some blurring.

Segmented PCT (SPCT) is a compromise between band selection and conventional feature extraction methods. The computational burden of SPCA is not heavy. It also contains the major data variance information and keeps the band information without losing interpretability. Although CPCA and ICA are more appealing theoretically, they did not show good performance in our case. CPCA emphasizes the YST classes, but the visual effects of the generated images suggest it does not reduce the variance of other classes. The behavior of ICA is similar to PCA and the fourth independent component channel shows a similar pattern as the first PCA channel. Therefore, the combination of the fourth ICA channel and the first and second LDA channels provides a similar visualization effect as shown in Plate 1c.

To judge the visual effect of the generated images, we have consulted the experienced field researchers and used field survey data as well as classification maps. We proposed three indicators, Infodex, Sepadex, and ACC to measure the information capacity, between-class separability and information redundancy of the generated images. We believe the image with enhanced visual effect should be a good trade-off among these three indicators. The images derived from SPCT and images combining one PCA channel with the positive and negative of the first LDA channels are both good examples. The former emphasizes more on the information capacity, and the latter emphasizes more on the between-class separability.

However, it is important to note that image visualization techniques cannot replace conventional land cover classification. The live YST patches cannot be visually separated from riparian vegetation in Plate 1c; over-mature YST seems to be overestimated on the west of the creek as shown in Plate 1e. The complex patterns of color in other composed images, for example, SPCT images as shown in Plate 1d and 1f, can still possibly hamper the intuitional understanding of the distribution of YST. Nevertheless, the visualization techniques provide an easy way to interpret the image at the early stage of image analysis. The preliminary information on the distribution of YST can provide a general guide for the further field survey.

Finally, more research is required to effectively extend our proposed visualization techniques to other hyperspectral applications. We need to test these algorithms for other multispectral and hyperspectral images with different spatial resolutions, or broader spectral ranges, or with larger spatial extents. Furthermore, evaluation of visual effects on subtle variation is a process that is difficult to obtain objective opinions. Although we have defined three indicators to describe the information capacity (Infodex), between-class separability (Sepadex), and information redundancy (ACC), they reflect only three aspects of the image quality and the trade-offs among them also requires subjective judgment. In this study, we tested the seven band selection and five feature extraction methods. Other dimension reduction algorithms based on different criteria need to be explored for hyperspectral image visualization.

Acknowledgments

The authors are grateful to the three anonymous references whose insightful comments were very useful for improving this paper.

References

Beaudemin, M., and K.B. Fung, 2001. On statistical band selection for image visualization, *Photogrammetric Engineering & Remote Sensing*, 67(5):571–574.

Bellec, B., and H. Legleau, 1992. The multispectral color-composite technique, An improved method to display meteorological satellite imagery, *International Journal of Remote Sensing*, 13(11):1981–1998.

Cardoso, J.-F., 1994. On the performance of orthogonal source separation algorithms, *Proceedings of EUSIPCO*, Edinburgh, Scotland.

Cardoso, J.F., and A. Souloumiac, 1993. Blind beamforming for non-gaussian signals, *IEEE Proceedings-F Radar and Signal Processing*, 140(6):362–370.

Chang, C.I., S.S. Chiang, J.A. Smith, and I.W. Ginsberg, 2002. Linear spectral random mixture analysis for hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 40(2):375–392.

Clevers, J., and H.T.C. Vanstokkom, 1992. The quantitative-evaluation of false color-photography with application of a red filter, *International Journal of Remote Sensing*, 13(9):1709–1733.

DiTomaso, J.M., and J. Gerlach, 2000. Centaurea solstitialis (yellow starthistle), *Invasive Plants of California's Wildlands*, (C. Bossard, J.M. Randall and M. Hoshovsky, editors), Berkeley, California, U.C. Press, pp. 101–106.

Duchene, J., and S. Leclercq, 1988. An optimal transformation for discriminant and principal component analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6):978–983.

Gong, P., 1993. Change detection using principal component analysis and fuzzy set theory, *Canadian Journal of Remote Sensing*, 19(1):22–29.

Gong, P., R. Pu, and B. Yu, 2001. Conifer species recognition: Effects of data transformation, *International Journal of Remote Sensing*, 22(17):3471–3481.

Haack, B., and S. Jampoler, 1995. Color composite comparisons for agricultural assessments, *International Journal of Remote Sensing*, 16(9):1589–1598.

Hsu, P.-H., 2002. Data exploration and analysis of hyperspectral images: Visualization and symbolic description, *Proceedings of the Asian Conference on Remote Sensing*, Kathmandu, Nepal.

Jia, X., and J.A. Richards, 1999. Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification, *IEEE Transactions on Geoscience and Remote Sensing*, 37(1):538–542.

Kailath, T., 1967. Divergence and Bhattacharyya distance measures in signal selection, *IEEE Transactions on Communication Technology*, CO15(1):52–60.

Landgrebe, D., 2002. Hyperspectral image data analysis, *IEEE Signal Processing Magazine*, 19(1):17–28.

Lass, L.W., H.W. Carson, and R.H. Callihan, 1996. Detection of yellow starthistle (Centaurea solstitialis) and common St. Johnswort (Hypericum perforatum) with multispectral digital imagery, *Weed Technology*, 10(3):466–474.

Lass, L.W., B. Shafii, W.J. Price, and D.C. Thill, 2000. Assessing agreement in multispectral images of yellow starthistle (Centaurea solstitialis) with ground truth data using a Bayesian methodology, *Weed Technology*, 14(3):539–544.

Lass, L.W., D.C. Thill, B. Shafii, and T.S. Prather, 2002. Detecting spotted knapweed (Centaurea maculosa) with hyperspectral remote sensing technology, *Weed Technology*, 16(2):426–432.

Mausel, P.W., W.J. Kramber, and J.K. Lee, 1990. Optimum band selection for supervised classification of multispectral data, *Photogrammetric Engineering & Remote Sensing*, 56(1):55–60.

Miao, X., P. Gong, R. Pu, R.I. Carruthers, and J.S. Heaton, 2007. Applying class-based feature extraction approaches for supervised classification of hyperspectral imagery, *Canadian Journal of Remote Sensing*, 33(3).

Milton, E.J., and K.Y. Choi, 2004. Estimating the spectral response function of the CASI-2, *RSPSoc2004: Mapping and Resources Management*, Annual Conference of the Remote Sensing and Photogrammetry Society, Aberdeen, Scotland.

Mooney, H.A., and J.A. Drake, 1987. The ecology of biological invasions, *Environment*, 29(5):10–20.

Okada, T., and S. Tomita, 1985. An optimal orthonormal system for discriminant-analysis, *Pattern Recognition*, 18(2):139–144.

Polder, G., and G.W.A.M. van der Heijden, 2001. Visualization of spectral images. *Proceedings of SPIE*, 23–24 October, Vol. 4553.

Pu, R., and P. Gong, 2004. Wavelet transform applied to EO-1 hyperspectral data for forest LAI and crown closure mapping, *Remote Sensing of Environment*, 91:212–224.

Pu, R.L., P. Gong, G.S. Biging, and M. Larrieu, 2003. Extraction of red edge optical parameters from Hyperion data for estimation of forest leaf area index, *IEEE Transactions on Geoscience and Remote Sensing*, 41(4):916–921.

Richards, J.A., 1993. *Remote Sensing Digital Image Analysis: An Introduction*, New York, Springer-Verlag.

Underwood, E., S. Ustin, and D. DiPietro, 2003. Mapping nonnative plants using hyperspectral imagery, *Remote Sensing of Environment*, 86(2):150–161.

Ustin, S.L., G. Scheer, D. DiPietro, E. Underwood, and K. Olmstead, 2001. Hyperspectral remote sensing for invasive species detection and mapping, *Abstracts of Papers of the American Chemical Society*, Vol. 221.

Vitousek, P.M., 1990. Biological invasions and ecosystem processes – Towards an integration of population biology and ecosystem studies, *Oikos*, 57(1):7–13.

Webb, A., 2002. *Statistical Pattern Recognition*, West Sussex, England, John Wiley and Sons.

Wilcove, D.S., D. Rothstein, J. Dubow, A. Phillips, and E. Losos, 1998. Quantifying threats to imperiled species in the United States, *Bioscience*, 48(8):607–615.

Xu, B., and P. Gong, 2005. Land use/cover classification with multispectral and hyperspectral EO-1 data, *Proceedings of the ASPRS Annual Conference*, Anchorage, Alaska, 05–09 May, unpaginated CD-ROM.

Yu, B., I.M. Ostland, P. Gong, and R. Pu, 1999. Penalized discriminant analysis of in situ hyperspectral data for conifer species recognition, *IEEE Transactions On Geoscience and Remote Sensing*, 37(5):2569–2577.

(Received 12 October 2005; accepted 1 November 2005; revised 31 January 2006)