Interest Points for Hyperspectral Image Data

Amit Mukherjee, Miguel Velez-Reyes, *Senior Member, IEEE*, and Badrinath Roysam, *Senior Member, IEEE*

Abstract—**Interest points are widely used as point-features for image matching. This paper describes robust and efficient algorithms to extract multiscale interest points in hyperspectral images in which structural information is distributed across several spectral bands. The formulation is based on a Gaussian scale-space representation of the hyperspectral data cube, and the use of a principal components decomposition to combine information efficiently across spectral bands. A spectral distance measure is used to characterize spatial relations between neighboring hyperspectral pixels. In addition, we describe methods for preprocessing a pair of hyperspectral images, clustering the spectral signatures of interest points, and using the resulting data for matching points under simple geometric transformations. The stability of the resulting interest points in time-lapse satellite images was determined to be in the range of 52% to 75% in the testing data set that were acquired from variety of landforms like coastal islands of La Parguera, Chesapeake Bay, the Cuprite Mining District of Nevada, and agricultural field images of Kansas and Oklahoma, and thus, they can be used as a foundation for image matching and related image analysis tasks.**

Index Terms—**Clustering, difference of Gaussian (DoG), hyperspectral image, interest points, key points, principal components analysis (PCA), scale-invariant feature transform (SIFT).**

I. INTRODUCTION

THE GOAL of this paper is to develop a multiscale interest point operator for hyperspectral image data. Interest point operators for monochromatic images were proposed a decade ago [1] and have been studied since then. They have been applied to diverse problems in image matching, recognition, registration [2], change detection, and content-based image retrieval. An interest point operator seeks out points in an image that are structurally distinct, invariant to imaging conditions, and certain expected geometric transformations [1], [3], [4]. Interest points are also helpful as a data reduction method in registration and object detection problems. The advantage of using interest points is to replace an exhaustive search over the entire image space by a computation over a concise set of highly informative points. Our goal in this paper is to extend the notion of interest points to hyperspectral image data.

Hyperspectral imaging, also known as imaging spectroscopy, is growing in importance as a method for spatially resolving classes of substances in diverse applications ranging from Earth surveillance to biomedical imaging [5]. Each pixel of a hyperspectral image is a uniformly sampled spectrum of the corresponding point in the scene [5]–[7]. The spectral resolution of modern hyperspectral imaging systems that operate in the reflective region of the spectrum is approximately 10 nm. Sensors like Hyperion scan the spectrum from about 400–2500 nm in intervals of 10 nm [8]. This resolution is high enough to allow the measurements at each to be treated as a continuous spectrum. The higher spectral resolution of hyperspectral images compared to multispectral images offers the potential to increase material discrimination and target detection accuracy. For instance, kaolinate and dickite have been shown to be distinguishable in hyperspectral imagery, but not in seven-band Landsat data because of the high similarity of their spectral signatures [5].

Several challenges must be overcome in order to achieve successful analysis of hyperspectral image cubes. One obvious challenge is the sheer size of data cubes, implying large computational requirements. For instance, images from Hyperion typically consist of 100–200 bands per pixel, and each band is encoded by a 16-bit quantity. Often, several of the bands (typically 20–40) are often unusable due to dead sensor elements, saturation/cutoff, or calibration errors [8]. Even after unusable bands are eliminated, image analysis algorithms must be designed to be robust to variations in illumination, atmospheric conditions, cloud cover, wave/tidal conditions, and seasonal/diurnal changes in reflectance. Each pixel in a hyperspectral image typically covers a large spatial region on Earth's surface (e.g., Hyperion has a 7.5 km swath width and a 30-m spatial resolution), and the pixel spectra are impure, i.e., they represent a complex mixture of end members [8]. In addition, images acquired using swept linear arrays (pushbroom systems) contain linelike scanning artifacts that exist across spectral channels. These artifacts create potential ambiguities for algorithms intended to extract structures in the images. Finally, varying degrees of imaging noise are present across the spectral channels. In summary, successful analysis of hyperspectral data requires image analysis methods that are robust to the expected imaging deficiencies while being computationally efficient.

In this paper, we formulated an interest point operator for hyperspectral data. Interest points are a subset of the structural features in an image such as edges, corners, and blobs. In hyperspectral images, these structural features are not contained in a single plane, but rather, distributed across multiple spectral channels. For example, an edge between two adjacent regions may be discernible in a band of the spectrum where the spectral

contrast is significant and undetectable elsewhere. Moreover, it is not possible to predict a spectral band in which all the structural information across the entire hyperspectral image is sufficiently prominent for reliable detection. One obvious approach to this problem would be to compute a linear projection of the hyperspectral cube to form a monochromatic image, and then use a conventional (gray scale/monochromatic) operator to detect structural features. A good choice of such a projection plane could be from principal components analysis (PCA) of the hyperspectral data. The PCA would "rotate" the hyperspectral data in such a way so that most of the variance in the data is spanned by the first few principal components. However, such an approach is unavoidably biased by the choice of the PCA components, and the spatial neighborhood over which PCA is performed. In this paper, we addressed this problem of detecting structural features that are distributed across several planes by combining image information across the spectral channels in a nonlinear manner.

The concept of interest points is well developed for monochromatic images like the gray scale images. The Harris operator [2], [9] and its many variations (e.g., see [4], [10], and [11]) have found widespread applications in computer vision and image processing since it was first proposed in 1988. Since then, several improvements have been described in the literature. The more significant improvements include Lowe's scale-invariant feature transform (SIFT) [3], Lindeberg's blob detection approach [12], Mikolajczyk's affine-invariant Harris–Laplace operator [13], Kadir's salient point detection [15], and some others [11]. It is desirable for an interest point to be invariant to the expected range of spatial transformations (translation, rotation, scaling), illumination changes, and limited changes in view point. The methods proposed by Lowe and Lindeberg require convolution of the image by a Gaussian kernel at multiple scales, and identifying the locations of local extrema (maxima and minima) representing bloblike structures. These local extrema are invariant to rotation, scale, and viewpoint changes to a limited extent [3], [4], [10] and, hence, form good interest points.

Some prior attempts have been made to extend interest point operators to three-color images, mainly normalized red/green/blue (RGB) images [17] acquired by conventional color cameras. RGB images can be thought of as a special low-dimensional case of multi/hyperspectral images. For instance, the C-SIFT algorithm [18] extends Lowe's keypoint operator to RGB images. This method relies on mathematical modeling of the reflected spectrum of objects in the visible band, and the response of the human visual system. For this paper, the C-SIFT strategy was not readily usable, although several concepts underlying Lowe's keypoint operator were directly applicable. Specifically, we used the difference of Gaussian (DoG) kernels for each spectral layer of the preprocessed hyperspectral image, following Lowe's approach. We also employed PCA to reduce the computational burden of scale-space decomposition of the multiscale DoG operator, and to achieve robustness to imaging artifacts that are not correlated with the data, e.g., line artifacts from pushbroom sensors, and faulty sensor elements.

In formulating the proposed method, we considered potential alternatives. For example, the Beltrami framework is an approach to multispectral image analysis based on treating the spatial and intensity dimensions in a uniform manner using the concept of a structure tensor [19]. A corner detector based on the eigenvalues of the structure tensor in the Beltrami framework would sum the square of image gradients in each spectral band to compute the overall structure tensor. Notwithstanding the elegance of the idea, it is not attractive for analyzing hyperspectral data since it is sensitive to noise and spatial irregularities that are commonly present in some channels of hyperspectral images. It is also computationally prohibitive, particularly considering the high dimension of a typical hyperspectral image.

II. OVERVIEW

Our approach to formulating an interest point algorithm for hyperspectral images was motivated by the observation that the natural images of interest to us lack strong corners. This has driven us to focus on region-based methods rather than edge-based methods [3], [4]. Specifically, we build upon methods that use local extrema (maxima/minima) of the results of filtering the image data by a DoG kernels. This filter approximates the scale-normalized Laplacian proposed by Lindeberg [12]. The filtering operation is understood most effectively by treating the image as a quasi-thermal field $U(\mathbf{x})$ that evolves in accordance with the heat diffusion equation. Starting from an initial condition $U(\mathbf{x}, t = 0)$, the field evolves according to the partial differential equation $U_t(\mathbf{x}, t) = 1/2U_{xx}(\mathbf{x}, t)$, where the subscripts indicate partial derivatives. The unique solution to this equation is given by $U(\mathbf{x}, t) = U(\mathbf{x}, 0) * g(t)$, where $g(t)$ is a Gaussian with scale t, and $*$ represents the linear convolution operation. Lowe showed that when the diffusion equation is approximated using finite elements, and $\sigma^2 = t$, it can be written as follows:

$$U(\mathbf{x}, k\sigma) - U(\mathbf{x}, \sigma) = (k - 1)\sigma^2 U_{xx}$$

where σ^2 is the variance of the Gaussian kernel, and k is the ratio between discrete scale steps [3]. Note that the left-hand side of the above equation represents a DoG kernels, whereas the right-hand side is proportional to the scale-normalized Laplacian proposed by Lindeberg [12]. Thus, the DoG operation approximates the scale-normalized Laplacian. The local extrema (local maxima and minima) of this solution along with some interpolation and threshold can be used to detect interest points.

Let $\mathbf{h}(\mathbf{x})$ denote a pixel vector of the hyperspectral image with L spectral bands, where $\mathbf{x} = [x, y]$ denotes the spatial coordinates. While referring to this pixel in scale space, we used the notation $\mathbf{h}(\mathbf{x}, \sigma)$, where σ denotes the scale. When the need arises to refer to a specific spectral band l, we used the notation $h(\mathbf{x}, \sigma, l)$. In defining an interest point for hyperspectral images, we were primarily interested in identifying spatial locations, and their spectral signatures. In other words, localization along the spectral dimension was not of interest. Accordingly, we considered diffusion processes independently in each spectral layer. Specifically, we considered DoG kernels at various scales applied to each spectral band of the hyperspectral image. Let

$\text{DoG}(\mathbf{x}, \sigma, l)$ denote the response of the DoG kernel at location $\mathbf{x}$, scale σ and spectral band l. This cylindrical kernel can be thought of as a stack of 2-D Gaussian kernels, one for each spectral band. As proposed by Lowe [3] and Burt [20], the scale of the DOG kernels were sampled in discrete steps, so that each octave was divided into an integer number of intervals, keeping the ratios of consecutive scales $k = \sigma_{i+1}/\sigma_i$, a constant.

In order to combine the $\text{DoG}(\mathbf{x}, \sigma, l)$ responses across N spectral bands, we proposed a nonlinear combining function $f_l(\bullet)$ that yielded a quantity $D(\mathbf{x}, \sigma) = f_l(\text{DoG}(\mathbf{x}, \sigma, l))$. The function $f_l(\bullet)$ is of the form $\mathbb{R}^L \rightarrow \mathbb{R}^+$, where $\mathbb{R}^+$ denotes the positive real axis. If the spatial neighborhood of $\mathbf{x}$ is denoted $\Omega(\mathbf{x}, \sigma)$, then the local maxima of $D(\mathbf{x}, \sigma)$ in the neighborhood, i.e., $\arg\max_{x, \sigma \in \Omega(\mathbf{x}, \sigma)}\{D(\mathbf{x}, \sigma)\}$ are the locations and scales of interest points. In later sections, we discuss some desirable properties of the function $f_l(\bullet)$ and compare a few of them by evaluating the performance of interest points for registering pairs of time-lapse hyperspectral images.

We begin by describing the application of PCA to make the diffusion computations efficient in Section III-A. In Section III-B, we describe a method to select a relevant subset of principal component projections. The scale-space representation and interest point detection methods, along with the associated subscale and subpixel interpolation methods are explained in Section III-C. Section III-D focuses on a few functions used in Section III-C to combine the information across spectral bands. The steps for detection of interest points are also summarized in Fig. 1. Section IV describes an implementation of interest point where the interest points of pairs of time-lapse hyperspectral images are evaluated by computing their repeatability in location and scale. Finally, we describe methods to reduce the sparseness of the match using hierarchical clustering of the spectral descriptors associated with interest points.

III. HYPERSPECTRAL INTEREST POINT DETECTION METHOD

A. Scale-Space Representation Using Principal Components

This section describes the use of principal component analysis (PCA). The use of PCA provides two major reasons: 1) fast implementation of DoG kernel computations on hyperspectral data and 2) reduction of unstable interest points arising from spatial irregularities and noise.

Hyperspectral images commonly contain a significant amount of redundancy. Thus, abstraction of information content from high-dimensional data forms one of the basic challenges in interest point detection. Satellite images typically cover large portions of Earth's surface (e.g., 30 m/pixel for Hyperion data), so random irregularities and variations in the data are not uncommon. In addition, one must consider factors such as noise [6], variations in illumination, and distortions in the image caused by atmospheric variations in time-lapse data sets. These factors have been studied by several researchers, and used in formulating the spectral radiance model [6]. From the standpoint of extracting interest points, these factors can increase the false detection rate unacceptably.

Classical PCA [22] seeks a linear decomposition of the data in a least squares sense. The PCA projection uses eigen-
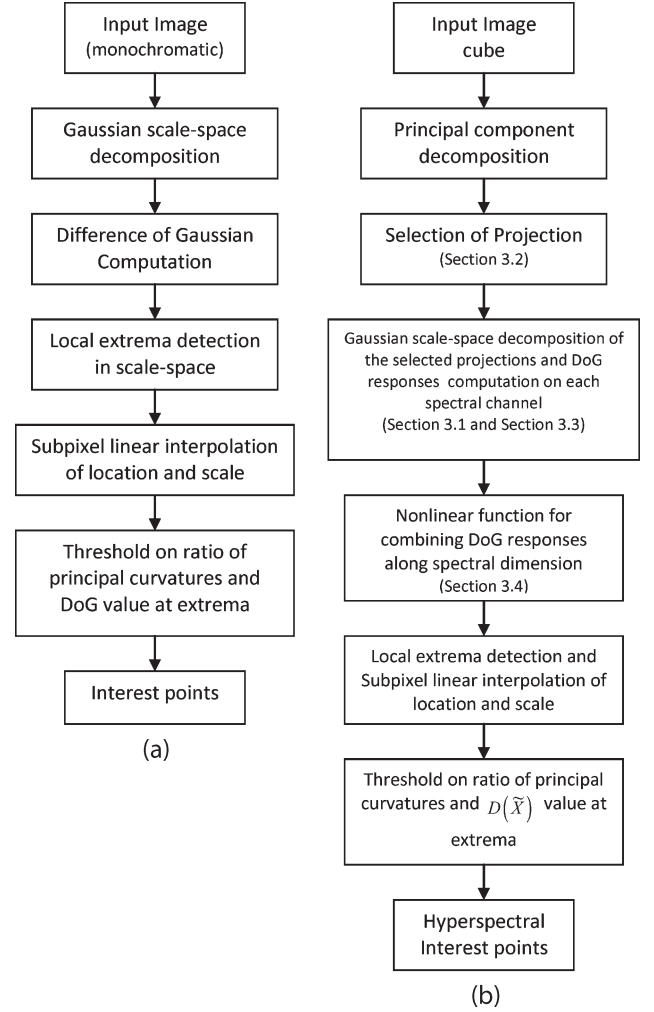
Fig. 1. (a) Flowchart showing various steps of Lowe's keypoint detection. (b) Flowchart showing various steps of hyperspectral interest point detection.

decomposition of the covariance matrix to form projection directions into orthogonal subspaces. The decomposition of a hyperspectral pixel along the spectral dimension is a linear transformation of the form, where denotes the value of the pixel vector when projected on the principal axis, is the mean, is the eigenvalue, and is the eigenvector, respectively. Typically, more than 99% of the variance is spanned by the top 5% of projection directions. Thus, PCA could potentially yield a roughly 20-fold data reduction. This redundancy is exploited by our algorithm to reduce the computations.

The responses of the DoG kernel for spatial location $\mathbf{x}$ and scale σ can be represented as follows:

$$D(\mathbf{x}, \sigma) = f_l\left(\text{DoG}(\mathbf{x}, \sigma, l)\right)$$

$$= f_l\left(\mathbf{h}(\mathbf{x}) \otimes g(k\sigma) - \mathbf{h}(\mathbf{x}) \otimes g(\sigma)\right) \quad (1)$$

where $g(\sigma)$ is the 2-D Gaussian kernel that is stacked L times (one for each band of the hyperspectral pixel) and $\otimes$ is a 2-D spatial-domain convolution for each spectral band. Decomposing $\mathbf{h}(\mathbf{x}) = \sum_i \alpha_i(\mathbf{x})\lambda_i \mathbf{v}_i$ using PCA and using the

distributivity of convolution over additions (see the Appendix), (1) can be simplified as follows:

$$D(\mathbf{x}, \sigma) = f_l \left(\left(\sum_i \alpha_i(\mathbf{x}) \lambda_i \mathbf{v}_i \right) \otimes g(k\sigma) \right.$$
$$\left. - \left(\sum_i \alpha_i(\mathbf{x}) \lambda_i \mathbf{v}_i \right) \otimes g(\sigma) \right)$$
$$= f_l \left(\sum_i (\alpha_i(\mathbf{x}) \otimes g(k\sigma) - \alpha_i(\mathbf{x}) \otimes g(\sigma)) \lambda_i \mathbf{v}_i \right)$$
$$= f_l \left(\sum_i \text{DoG}(\alpha_i(\mathbf{x}), \sigma, i) \lambda_i \mathbf{v}_i \right). \tag{2}$$

Note that in the above derivation, only $\alpha_i(\mathbf{x})$ and $g(\sigma)$ are functions in space. The remaining quantities can be treated as scalars for a specific spectral band l, and accordingly, taken out of the convolution operation. For simplifying the notation, let $\mathbf{y}(\mathbf{x}, \sigma) = \sum_i \text{DoG}(\alpha_i(\mathbf{x}), \sigma, i) \lambda_i \mathbf{v}_i$, so that we can denote the function $D(\mathbf{x}, \sigma) = f_l(\mathbf{y}(\mathbf{x}, \sigma))$, where $\mathbf{y}$ is a vector of length N containing the DoG values for each spectral band in the hyperspectral image.

In the above equation, we have separated the term $\lambda_i \mathbf{v}_i$ from $\alpha_i \in \mathrm{A}_i$, the projected values on the axis for the ith principal component. Therefore, we have shown that the DoG kernel can be computed from the projections on the principal component axes instead of the hyperspectral image.

B. Projection Selection

In this section, we show that by thoughtfully selecting a subset of projections from the overall eigenspace, we can address two issues simultaneously: 1) computational efficiency and 2) filtering of spatial irregularities for extracting stable interest points. As noted earlier, there is a high degree of correlation among the spectral signatures of hyperspectral pixel vectors, so a modest number of principal components can capture most of the spectral information [23]. Our goal in this section is to select the components that: 1) yield a stable set of interest points and 2) minimize the computational burden.

A common method to choose an effective subset of principal components is to retain only the components with the highest eigenvalues. Eigenvalues estimate the spread (variance) of the data in the direction of the corresponding eigenvectors, and therefore, their magnitudes indicate the amount of structural information spanned by the corresponding projections. The disadvantage of this method is the difficulty in choosing an automated threshold on the eigenvalues, so that the principal components below that threshold can be discarded for further processing. In addition, we wanted to retain as much spectral information as possible, as long as the projection subspace is acceptably smooth (i.e., with the fewest spatial irregularities). With this in mind, we proposed the following strategy: Initially, any principal components that span more than 5% of the data variance are retained. For all the remaining principal components, we retained those principal components whose projections are acceptably smooth in the sense defined below.

As a metric to quantify the smoothness of a component projection A_i, we proposed the following "smoothness ratio":

$$S_{n \times n} = std(\hat{\mathrm{A}}_i)/std(\mathrm{A}_i - \hat{\mathrm{A}}_i)$$

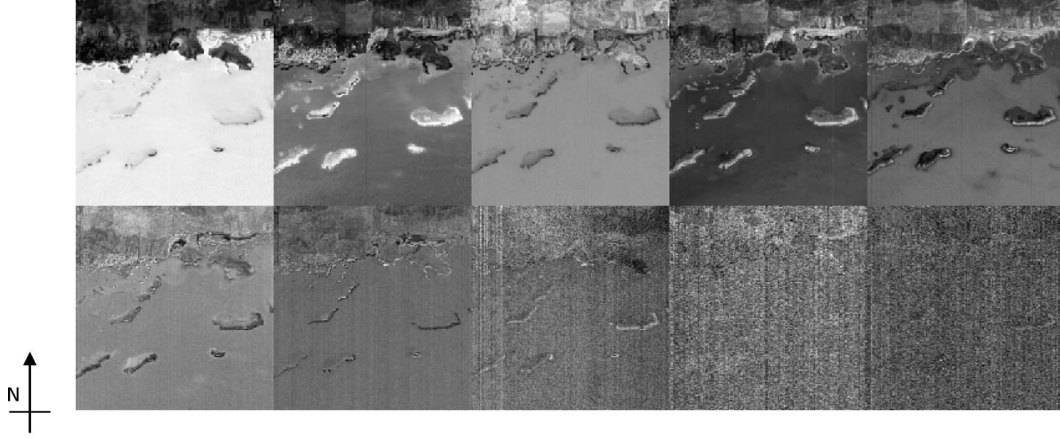where $\hat{\mathrm{A}}_i$ is the minimum mean absolute error estimate in the local $n \times n$ neighborhood of each pixel and it is obtained by median filtering, and $std(\bullet)$ refers to the standard deviation. Projection bases for which this ratio is greater than one are labeled as smooth projections, and added to the list of projections with higher eigenvalues. Intuitively the "smoothness ratio" compares the standard deviation of the smoothed image $(\hat{\mathrm{A}}_i)$ to the standard deviation of high-frequency components in the image that are obtained by subtracting the smoothed image from the original image $(\mathrm{A}_i - \hat{\mathrm{A}}_i)$. We used a median filter to perform the smoothing. When this ratio was greater than one, the smoothed image had higher standard deviation than the high-frequency components. In this case, we selected the projection for further processing. However, when the high-frequency components were more dominant, the smoothness ratio became less than one. These projections were not included for further processing because they gave rise to unstable interest points. The median filter is attractive because the filter response is robust to the presence of edges and impulsive outliers in the projection.

Fig. 2 shows a few of the principal projections of a sample hyperspectral image of size 250×239 pixels, and 126 spectral bands. This is an image of the La Parguera region off the island of Puerto Rico. The panels show the ten principal projections with the highest standard deviation and the corresponding smoothness ratios. It is quite clear from this display that anomalous outliers are frequent in most of the lower variance projections. Note that the two criteria set for selecting projections are not exclusive, because most of the top projections also have smoothness ratios less than one in our testing data set. Moreover, note the increased presence of imaging artifacts in projections for which the smoothness ratio is less than or equal to one. For this illustration, a window width of $n = 5$ pixels was chosen empirically. Decreasing the window width to four or three would qualify more projections making the smoothness ratio less strict. Thus, using the projection selection, a subset of the projections containing significant structural information of the hyperspectral image was identified that resulted in eliminating weak irregularities making the detector robust and stable. This also eliminated the vertical line artifacts, often seen in line-scan images.

C. Local Extrema Detection

Let B denote a subset of projections selected using the method described in Section III-B. Then, the hyperspectral image reconstructed from this subset of projections is given by $\hat{\mathbf{h}}(\mathbf{x}) = \sum_{i \in B} \alpha_i(\mathbf{x}) \lambda_i \mathbf{v}_i$. The corresponding value of function $D(\mathbf{x}, \sigma)$ is given by

$$D(\mathbf{x}, \sigma) \approx f_l \left(\sum_{i \in B} (\alpha_i(\mathbf{x}) \otimes g(k\sigma) - \alpha_i(\mathbf{x}) \otimes g(\sigma)) \lambda_i \mathbf{v}_i \right)$$
$$\approx f_l \left(\sum_{i \in B} \text{DoG}(\alpha_i(\mathbf{x}), \sigma, i) \lambda_i \mathbf{v}_i \right). \tag{3}$$

PCA projection	Standard deviation	Smoothness ratio $S_{5\times 5}$
1	813.29	7.53
2	248.49	2.97
3	127.29	1.81
4	71.49	2.27
5	46.64	2.38
6	31.26	1.38
7	20.18	1.00
8	16.32	0.97
9	15.15	0.50
10	14.16	0.43

Fig. 2.   Top ten principal projections (sorted by the variance) and corresponding smoothness ratios of the Parguera 2002 hyperspectral image. The projections are linearly scaled such that white is maximum value black is the minimum for each pane. Projections one–ten are arranged first left to right first, then top to bottom. The North direction is approximately shown by arrow in the top left, the scale of the images can be obtained from the relation that one pixel in the image corresponds to 30 m × 30 m on land.

Note that $g(\sigma)$ in the above equations is a stack having the same number of Gaussian kernels as there are projections in subset B. We have not yet specified the function $f_l(\bullet)$ that has a form suitable to map the DoGs vector into a positive scalar quantity. This is discussed in Section III-D.

In this section, we briefly describe the remainder of the interest point detection algorithm that is mostly built upon the work of Brown and Lowe [3], [17]. First, the scale σ is discretized as follows:

$$\sigma_s \in \{k^s\}; \quad k = 2^{1/3}, \qquad s = \left\{-1, 0, 1, 2, 3 \ldots, \frac{\log(s_{\max})}{1/3 \log 2}\right\} \tag{4}$$

where the exponent three implies that each octave of the scale space is sampled at three points. The parameter $s_{\max}$ is specified by the user based on the maximum width (in pixels) or the scale of the interest points desired for a given image. Note that the minimum scale due to the discretization is 1 pixel. Then, for each $A_i \in B$, $\mathrm{DoG}(\alpha_i(\mathbf{x}), \sigma, i)$ is computed for all scales, and using function f_l, the function $D(\mathbf{x}, \sigma)$ is computed for each scale and location of the hyperspectral image.

The computation of $D(\mathbf{x}, \sigma)$ was followed by the detection of local extrema in the 3-D space corresponding to image space $\mathbf{x}$ and σ. Each pixel was compared to its 26 neighbors that include eight immediate neighbors from the same scale and nine neighbors from two adjacent scales on either sides of σ.

It was selected as a candidate interest point if the $D(\mathbf{x}, \sigma)$ value was greater than that of all its neighbors. This step was followed by accurate interpolation of scale-space, using the Taylor series expansion (up to second degree) of $D(\mathbf{x}, \sigma)$ in the neighborhood of $\mathbf{x}$ and σ. This operation gave the localized candidate interest points (before thresholding) as

$$D(\tilde{X}) = D(X - \tilde{X}) + (X - \tilde{X}) \frac{dD}{dX} - (X - \tilde{X})^{\mathrm{T}} \frac{d^2 D}{dX^2} (X - \tilde{X}). \tag{5}$$

where $X = [\mathbf{x}, \sigma]^{\mathrm{T}}$. If the spatial offset of the Taylor series expansion was greater than 0.5 pixels, it was recomputed at the new location until the offset was small, or discarded if a set maximum number of iteration was reached. Following Lowe's algorithm, we defined a quantity called the "ratio of principal curvatures," that is the ratio of the larger and smaller eigenvalues of the spatial 2-D Hessian at location $\tilde{X}$. Among the set of candidate interest points, those with the ratio of principal curvatures less than ten were retained as interest points. The interested reader is referred to the paper by Lowe [3] for further details. In the following section, we describe the nonlinear combining function f_l used in (3).

D. Choice of the Nonlinear Combining Function

The nonlinear combining function $f_l(\mathbf{y})$ is of the form $\mathbb{R}^L \to \mathbb{R}^+$, mapping the DoG values for each spectral band(which

approximates the Laplacian value at a particular scale) in the hyperspectral image into a scalar quantity $D(\mathbf{x}, \sigma)$. The scalar value of $D(\mathbf{x}, \sigma)$ was used for locating the extrema (as mentioned in Section III-C) indicating the presence of interest points. In searching for a suitable form of this function, we evaluated several potential alternatives, borrowing from the literature on histogram/spectral comparison: the L1-norm, L1-top50, earthmovers' distance (EMD) metric [24], and the diffusion distance metric [25]. These distance metrics were not used in their classical form, but were modified to suit our application, keeping the desirable properties that they posses. These modifications are described in the following paragraphs.

One desirable property of the combining function $f_l(\mathbf{y})$ is that it should be sensitive to small variations. Since we were sampling a continuum of scale space, the DoG values are expected to vary smoothly in $\mathbb{R}^{2 \times N}$ and space without discontinuities [3]. It is therefore desirable that f_l should be sensitive enough to capture small changes in the underlying field. As an illustrative example, consider the simple formulation $f_l(\mathbf{y}) = \sum_i y_i$. This function fails to identify the change in $\mathbf{y}$ if the increase in some of the summed components equals the decrease in others, so the sum remains constant (for example $\{1, 1, 1\}$ becomes $\{1, 2, -1\}$). Instead, had we considered the L1 norm $f_{\mathrm{L1}}(\mathbf{y}) = \sum_i |y_i|$, this change would have been detected. In this sense, the L1 norm is more sensitive than $f_l(\mathbf{y}) = \sum y$. A variation of the L1 norm is the L1-top50, defined as $f_{\mathrm{top50}}(\mathbf{y}) = \sum_{y_i \geq \mathrm{median}(|\mathbf{y}|)} |y_i|$. Whereas f_{L1} sums all the elements of $|y|$, f_{top50} just sums the 50 highest absolute values of y. The f_{top50} was considered because it serves as a robust version of f_{L1}, by ignoring the smaller changes.

A fundamental limitation of the above formulations is the fact that they ignore the spectral coordinates in the distribution. This motivated us to consider cross-bin distance measures such as the diffusion distance and the EMD. Both are sensitive to the location of the bins in the spectrum in addition to the counts at each bin, and therefore, they use the bin indices of the spectra in the hyperspectral image. Since the DoG vector is a difference vector, we will consider $\mathbf{y}$ as being equivalent to the vector difference between corresponding bins of the spectra.

The diffusion distance introduced by Ling in [25], is already defined in terms of spectral differences as $f_{\mathrm{diff}}(\mathbf{y}) = \sum_{i=1}^{P} d_i(\mathbf{y})$, where P is the length of the pyramid formed by convolution with a Gaussian filter followed by dyadic decimation, and $d_i(\mathbf{y})$ is L1-norm of $\mathbf{y}$ at ith level of the pyramid. $f_{\mathrm{diff}}(\mathbf{y})$ seeks differences in the frequency information of $\mathbf{y}$ in addition to the bin level differences, thus extending the definition of $\mathbf{y}$ from 1-D Euclidean space to 2-D Hilbert space. The multiscale formulation ensures that the coarser scales are weighted more heavily compared to the finer scales.

The EMD proposed by Rubner [24] is another cross-bin comparison measure. Following Rubner's terminology, the EMD estimates the average displacement of "mass" required to transform the first spectrum into the second. Consider the case when the "ground distance" along the spectral dimension between bins is in L1, and the "cost" of transporting mass between matched bins is zero. Then, the total "work done" can be written as, $W(\mathbf{y}, F) = \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{n} f_{ij} d_{ij}$, where f_{ij} is the positive optimal flow from ith bin in $\mathbf{y}_+$ to jth bin in $\mathbf{y}_-$. In the above

expression, d_{ij} is the normalized L1 distance between the ith and jth bins divided by the L1 distance from the first to the last bin. Moreover, $\mathbf{y}_+$ and $\mathbf{y}_-$ refer to two vectors formed by splitting $\mathbf{y}$ into two parts such that $\mathbf{y}_+ = |\mathbf{y}|$ if $y_i > 0$, and 0 otherwise. Similarly, $\mathbf{y}_- = |\mathbf{y}|$ if $y_i < 0$, and 0 otherwise. Since the sum of $\mathbf{y}_+$ and $\mathbf{y}_-$ must be equal, we added the remaining unmoved mass (following EMD terminology) which is equal to $|\sum \mathbf{y}|$ to W, and therefore our modified EMD is given by

$$f_{\mathrm{emd}}(\mathbf{y}) = W(\mathbf{y}, F) + \left| \sum \mathbf{y} \right|. \qquad (6)$$

Note that unlike the original EMD, the work done was not normalizes by the total flow. Thus, like the diffusion distance, EMD also captures the spatial bin relations; however, unlike the diffusion distance, it is void of any basis function. In comparison to the diffusion distance, the EMD performs a linear rather than dyadic sampling of the scale. A summary of combining functions is presented in Table I(a). The next section presents an empirical evaluation of the various $f_l(\bullet)$ formulations, along with a discussion of preprocessing issues specific to image registration.

IV. EVALUATION OF INTEREST POINTS FOR REGISTRATION

An important use of interest points is in feature-based image registration [2], [27]. In order to test the validity of the feature points, we used time-lapse hyperspectral images, and computed interest points on them. A good interest point detector is expected to capture features that are invariant to the geometric and illumination changes in the scene. And thus, the interest points detected in the first image should match to the corresponding location in the time-lapse image. This property is useful for feature-based registration. In this section, we demonstrate how the detected interest points can be used for image matching and feature-based registration on time-lapse hyperspectral images.

As a data preprocessing step, we proposed the following scheme to reduce illumination variability in the data pair. PCA was computed jointly between the pair of images to get the same basis vectors for projection. Then, for each of the projection bases, the 25th and 75th percentile distributions of the projected data for the second image was scaled to match the 25th and 75th percentile distribution of the projection of the first image. This scaling assumes that at least 50% of the image was overlapping in location. These scaled principal component projections were used further for detecting interest points.

The scale-space formulation and extrema detection were implemented in MATLAB. For fast convolution, we used radial Gaussian filters in the frequency domain. Four hyperspectral image data pairs were used in the evaluation. All the data pairs were time-lapse images acquired by the Hyperion sensor onboard the Earth Observing 1 (EO1) Satellite [8]. The first data set was recorded near the La Parguera island (a fishing village in southwestern Puerto Rico), and it was obtained from the Center for Subsurface Sensing and Imaging Systems (CenSSIS) image database [26]. The data contains 126 wavelength bands from 447.17 to 2324.91 nm. These time-lapse hyperspectral

TABLE I

(A) TABULATION OF THE NONLINEAR EMPIRICAL FUNCTIONS SHOWN IN SECTION III-D. (B) INTEREST POINT DETECTION AND MATCHING USING THE FOUR EMPIRICAL FUNCTIONS. THE TABLE ALSO SHOWS THE NUMBER OF MATCHING PAIRS IN SCALE AND LOCATION

(a)

L1 metric	$f_{L1}(\mathbf{y}) = \sum_i |y_i|$
L1 robust metric	$f_{top50}(\mathbf{y}) = \sum_{y_i \geq median(|\mathbf{y}|)} |y_i|$
Diffusion metric	$f_{diff}(\mathbf{y}) = \sum_{i=1}^{p} d_i(\mathbf{y})$ $d_i(\mathbf{y})$ is $|\mathbf{y}|$ at i^{th} pyramid level
Earthmovers metric	$f_{emd}(\mathbf{y}) = W(\mathbf{y}, F) + |\sum \mathbf{y}|$. Refer Equation (6)

(b)

	L1 metric f_{L1}	L1 robust metric f_{top50}	Diffusion metric f_{diff}	Earth Movers metric f_{emd}
Parguera 2002	967	916	983	1042
Parguera 2003	1088	1182	1094	1146
Matching pairs	660 (64.34%)	679 (65.25%)	686 (66.15%)	746 (68.26%)
Chesapeake Bay 2002	974	943	975	961
Chesapeake Bay 2004	953	947	974	977
Matching pairs	679 (70.47%)	689 (72.91%)	695 (71.31%)	683 (70.48%)
Kansas & Ohio Agr. field 2002	1091	1085	1099	1106
Kansas & Ohio Agr. field 2004	990	1009	1008	991
Matching pairs	548 (52.72%)	570 (54.47%)	562 (53.39%)	554 (52.91%)
Cuprite Mining 2002	1196	1194	1193	1188
Cuprite Mining 2004	1213	1226	1219	1218
Matching pairs	885 (73.47%)	893 (73.81%)	907 (75.21%)	888 (73.82%)

images were captured in 2002 and 2003, respectively. The remaining data were obtained from the EO1 website of the U.S. Geological Survey (http://eo1.usgs.gov). We selected small segments from three time-lapse image pairs from the database, Chesapeake Bay, the Cuprite Mining District of Nevada, and agricultural field images of Kansas and Oklahoma. Of the 220 spectral channels in each image, uncalibrated channels, channels having saturated data, or dead/ flat detectors were not used. The final images had between 157 and 183 spectral bands from 357 to 2576 nm. Since these images were large in size (> 3000 scan lines), random region comprising about 200 scan lines from the first image in a time-lapse pair and approximately overlapping region from the other image in the pair were cropped and used for our experiments.

The performance of interest point detection was evaluated based on the repeatability of the interest points between the time-lapse pair of hyperspectral images. Repeatability is defined in terms of matching position and scale of the interest points in the two images. Please note that we did not consider matching the orientations of the interest points as in the SIFT algorithm, because we have not defined the notion of orientation for hyperspectral interest points. We defined the matching criteria as follows: A pair of interest points from the images were said to be "matched" if their relative displacement was less than their average scale, and the ratio of their scales was between 0.5 and 2. The geometric transformations between the images were estimated by manually selecting identical points in the two images, and fitting an affine transformation model [27]. Most of the analysis in this paper was done based on the percentage of the matched interest points, which is the ratio of the number of matching pairs of interest points to the geometric mean of the number of interest points found in the overlapping area of the two images. In Figs. 4, 5, and 7, interest points are displayed as a marker, and the scale is indicated by an ellipse surrounding the interest point. The orientation, major axis, and minor axis of the ellipse were estimated from the Hessian of $D(\mathbf{x}, \sigma)$, at the location of the interest point.

Although a direct comparison with Lowe's SIFT [3] algorithm was not possible, we at first checked to see if the current algorithm showed any improvement over the SIFT in the hyperspectral image. Lowe's work has proven to be very successful over past several years for a variety of research efforts. However, SIFT was developed for single-layered (gray scale) images, and we could not use it directly for hyperspectral images. Therefore, we experimented by using SIFT on a pair of normalized projections of the hyperspectral images. We used the same PCA projections that were used in the interest point detection algorithm and examined the matching results. The results of matching the keypoints using Lowe's SIFT algorithm on the top ten PCA projections are shown in Fig. 3. The PCA was computed from the singular value decomposition of the combination of the all the pixel spectra of Parguera 2002 and Parguera 2003 images, similar to the hyperspectral interest point detection algorithm. Fig. 3(a) shows the percent correct matches out of the total number of interest points detected on the top ten PCA projections. Fig. 3(b) shows the total number of key-points detected from each projection. We observed that due to the high redundancy in the hyperspectral data, the first few PCA projections contained most of the structural information in the image, and were able to capture the matched interest points with high repeatability. It was also seen that the number of correct matches decreases significantly after the fifth PCA projection, because the data in these projections consisted mostly of noise and spectral variations that were not consistent across the two hyperspectral images. We also observed that the matching results were always below 50%—even for the dominant projection directions. The matching results reached a peak of 42.76% for PCA#2. This was partly because it is difficult to capture all the structural information by projecting the data onto a single plane. Moreover, this can only be achieved if we had prior knowledge of which PCA projection has to be
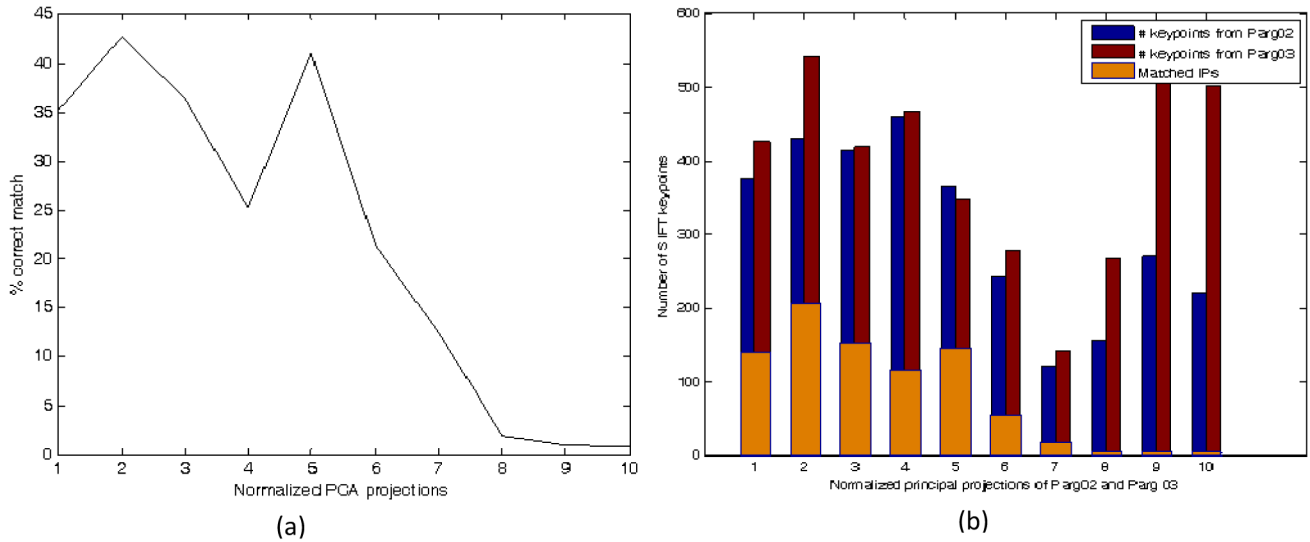
Fig. 3.   Result of matching the location and scale of the keypoints using the Lowe's SIFT operator on the top ten principal projections of the hyperspectral image. (a) Shows percentage of correct match when PCA projection is used in the top ten principal planes of Parguera 2002 and Parguera 2003. (b) Shows total number of keypoints detected in top ten principal projection planes along with the number of correct matches.

used. Without such knowledge, the task of matching becomes difficult. Thus, this experiment established the necessity of hyperspectral interest points by showing the limitations of Lowe's algorithm applied to PCA projections.

Next, we show the performance of our algorithm for interest point detection in hyperspectral image pairs. Table I(a) shows a summary of the combining functions $f_l(\bullet)$, and Table I(b) shows a quantitative evaluation of the interest point detector for different combining functions $f_l(\bullet)$ for the test data. It can be seen that all the empirical functions exceeded the maximum value of the matching performance measured on Lowe's keypoint detection algorithm shown in the previous paragraph. We also achieved approximately 20-fold reduction in computation cost of convolution by using PCA for Gaussian scale-space representation. For example, our projection selection criteria picked the top six principal components for Parguera 2002 image, and the top seven principal components for Parguera 2003 image out of the 126 eigenprojections. This step made significant reduction in the number of convolution operations for each scale at the expense of few additional multiplications and additions for combining the eigenvalues. We also observed that the proposed algorithm detected a greater number of interest points compared to what can be obtained from a single projection. This is due to the fact that the hyperspectral interest point detector uses information from the entire hyperspectral image and not from any selected spectral band.

Table I(b) shows that the matching percentage is high, the highest being 75.21% for the Cuprite mining district image pair using diffusion metric, and the lowest being Kansas and Ohio Agricultural field image pair using the L1 metric. Analysis of the total number of matching interest points showed that f_{diff} and f_{emd} yields more interest points than the other two. This revealed their higher sensitivity as compared to the first two functions.

Fig. 4(a) and (b) shows the interest points detected using the f_{diff} measure for the La Parguera data set magnified near the land-sea border. Fig. 4(c) and (d) shows the matching interest points between the two images. Similarly, Fig. 5(a)–(f) shows the interest points for the Cuprite mining district data set. Note that we have overlaid the hyperspectral interest points on an arbitrarily chosen principal projection to give readers a partial idea of the underlying datacube. The readers may not find the detection of some of interest points intuitive relative to the chosen projection. These interest points may be due structural information available in some other projections not visible in the chosen projection. For example, there are a few interest points on the seabed in the Parguera image in Fig. 4. However, since we have overlaid the interest points on PCA#2 (which does not show structures in the seabed), these interest points might appear counterintuitive if we try to relate the given image with interest points detected.

Due to the time-lapse nature of the image pairs, they are not exactly identical. We found several changes between the pair of images that are potential source of false alarm, like the presence of structures in the sea in the 2003 Parguera image, but not in 2002, and the presence of cloud cover in 2002 Chesapeake Bay image but not in 2004. In spite of these changes, our interest point detector proved to be robust showing satisfactory location and scale matches.

Another method to increase the confidence of matching interest points is by pruning the weaker responses D of the detector. This may result in a dense match because the strength of the response is an indication of its stability. Fig. 6(a) shows the effect of pruning the weakest interest points by thresholding. The weakest interest points can be eliminated by sorting the interpolated $\hat{D}(\tilde{\mathbf{x}}, \sigma)$ values and discarding the lowest n percentile of points. We have used $n = 30$ for our experiments. This leads to an increase in confidence of detection, as shown in the bar chart in Fig. 6(b).
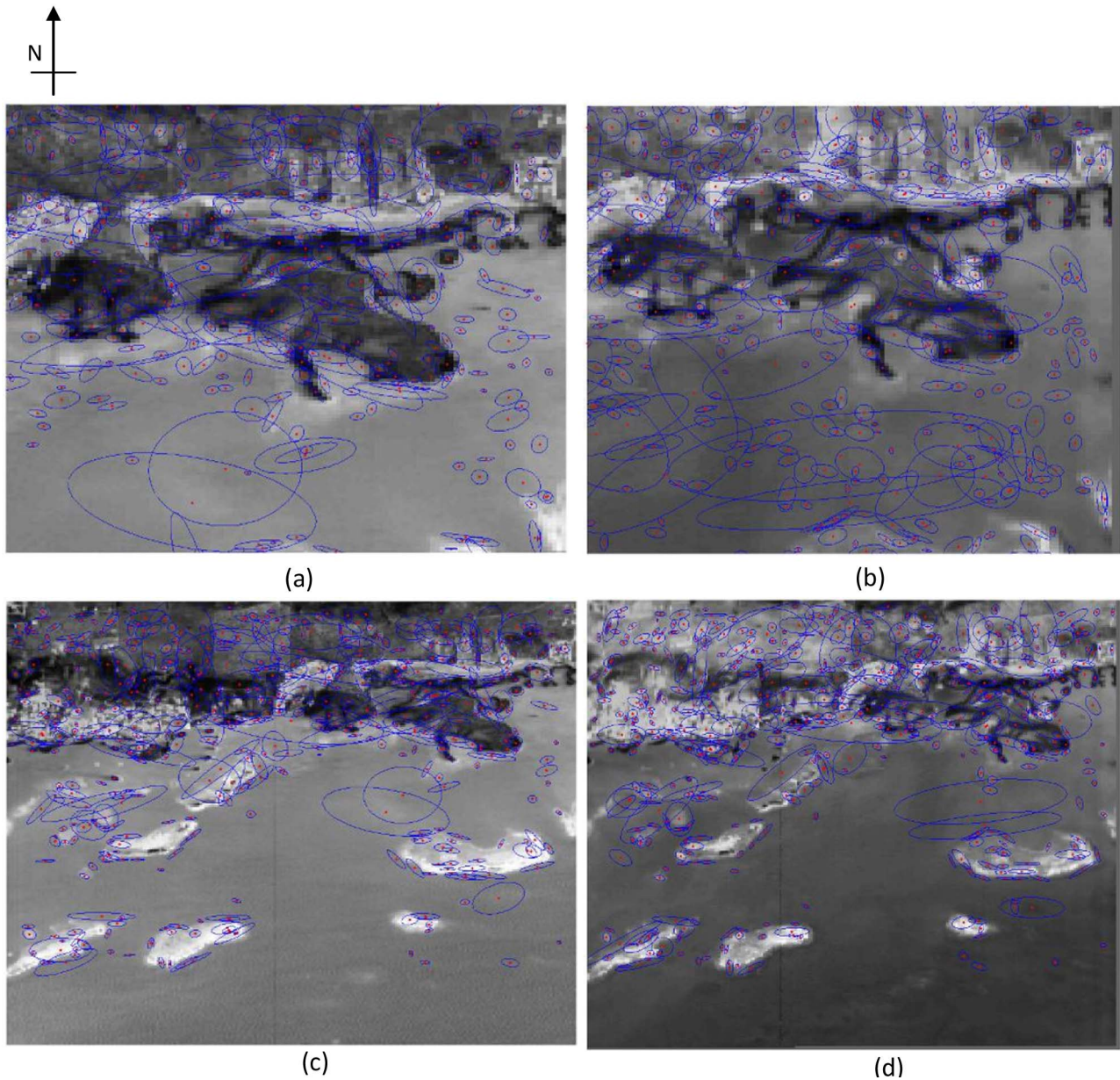
Fig. 4. Shows the detected interest points with using the diffusion metric with the affine scale shown by an ellipse surrounding the interest point. PC#2 of La Parguera Images are shown in background for reference. (a) and (c) shows image taken in 2002 whereas (b) and (d) shows image in 2003. The top panes, i.e., (a) and (b), shows a zoomed region of the image with all interest points, while the bottom pane, i.e., (c) and (d), shows the matched interest points in the entire image. One pixel in the image corresponds to 30 m × 30 m on Earth's surface.

We next show that a subset of interest points can be selected by clustering the average spectral signatures associated with the interest points. In the case of hyperspectral images, each pixel has a unique spectral signature associated with it, which is due to the combination of the spectral-reflectance of objects that are present in the space covered by the pixel location. We used the mean spectrum at the location and scale of an interest point as a feature for the interest points and clustered them hierarchically. The cluster centers were estimated using the mean spectrum of the members, and the Euclidean distance metric was used to form hierarchical clusters. A dendrogram presentation of the clustering results is shown in Fig. 6(b). The number of clusters was arbitrary selected as ten.

An analysis of clustering is shown in Fig. 7 and Table II. It can be seen that the interest points have formed natural clusters, for example, those in the sea, land, and coastal islands can be seen as distinct clusters. The clustering performed on interest points of the Parguera 02 image, and for each interest point in the clusters, matching interest points in scale and location is found in the Parguera 03 image. This is shown in columns 3 and 4 of Table II. It can be seen that for some clusters, e.g., Cluster 5, which belongs to the coastal region, the matching percentage is very high. However, for some other regions, e.g., Cluster 7 belonging to the sea, the matching rate is low.

Furthermore, the spectral signature of each matched interest point in Parguera 2003 is compared to the centers of the
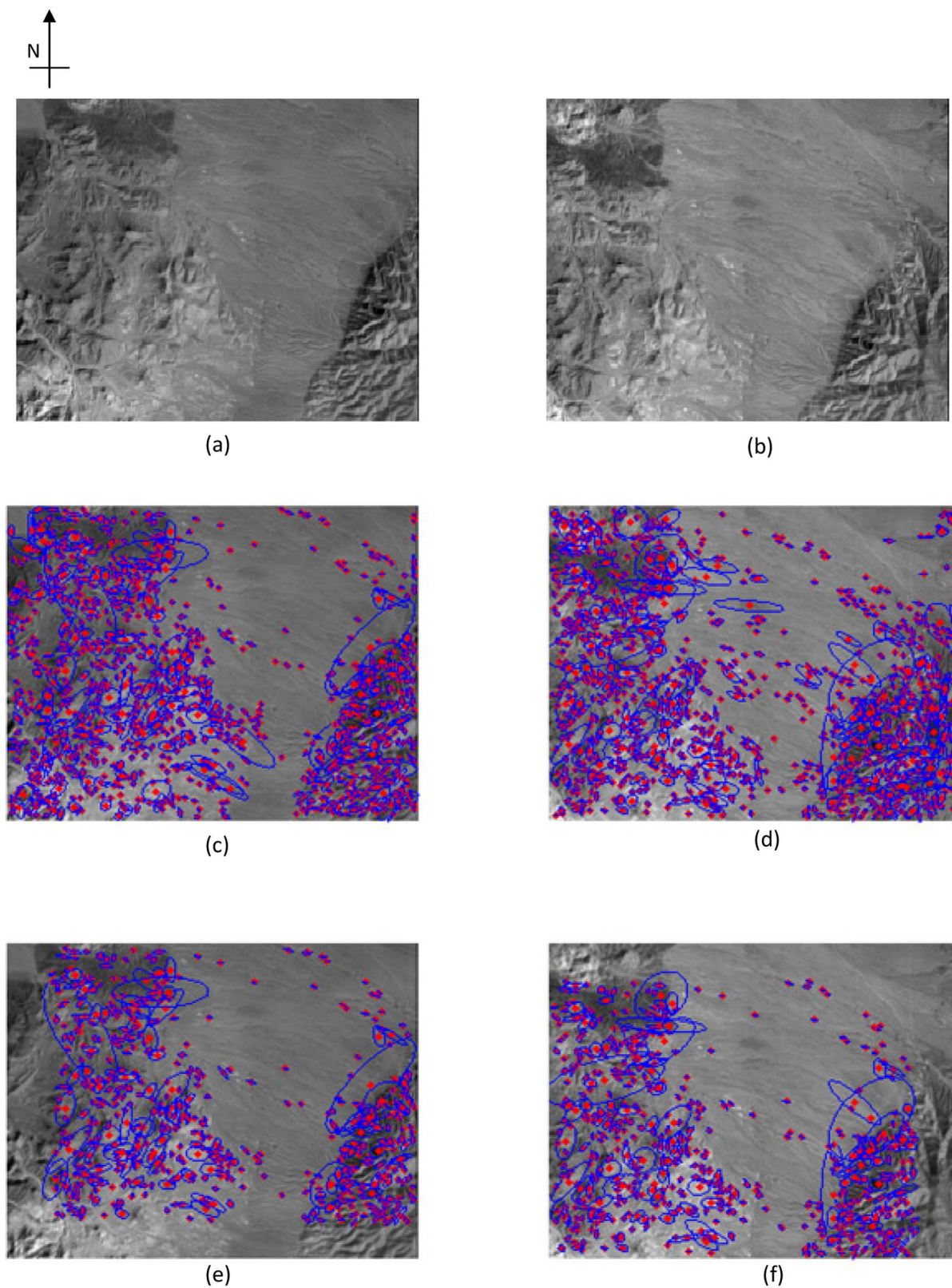
Fig. 5. (a) and (b) shows the PC#1 of the Cuprite mining districts hyperspectral image in 2002 and 2004, respectively. (c) and (d) shows the interest point detected in the image using earthmovers' metric as the nonlinear combining function. (e) and (f) shows the matching interest point in location and scale. The affine scale is shown by an ellipse surrounding the interest point. The scale of the images can be obtained from the relation that one pixel in the image corresponds to 30 m × 30 m on land.

clusters in Parguera 2002 using the Mahalanobis distance with the covariance estimated from the clusters in Parguera 2002 [Fig. 6(b)]. If the spectral signature lies within three standard

deviations, we denote that as a spectral match. Column 5 shows number of interest points matching in scale and location that qualifies the spectral match test. It can be seen that there is
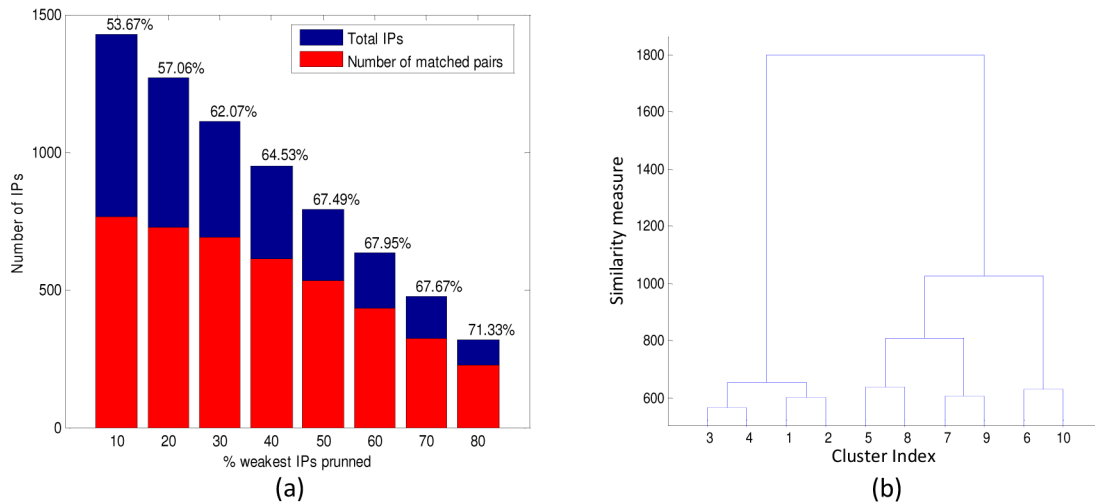
Fig. 6. (a) shows the effect of pruning the interest points based on the response of the DOG filter response. (b) shows the dendrogram constructed by hierarchical clustering of the spectral signatures of the interest points. The y-axis on the right is the similarity measure between clusters, and the x-axis shows the cluster IDs.
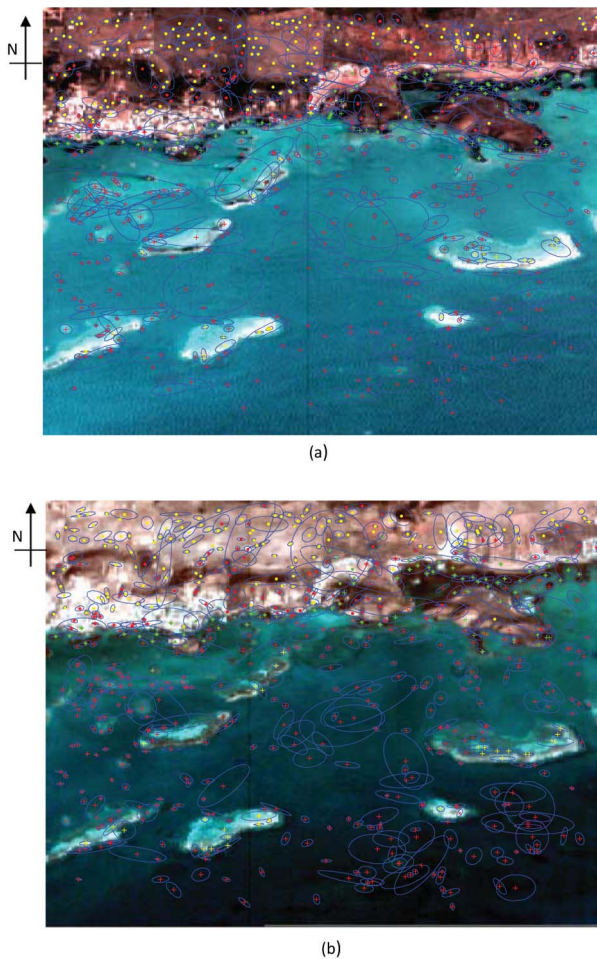

Fig. 7. La Parguera image pair with the interest point clusters shown in different colored markers overlaid on the RGB projections of the hyperspectral images. (a) shows interest points with markers representing clusters on the Parguera 2002 image. (b) shows interest points with markers representing clusters on the Parguera 2003 image. The ellipse around the interest points indicate the scale at which they were detected. Please refer to Table II for analysis of the interest points. The scale of images is as follows: one pixel in the image corresponds to 30 m × 30 m on Earth's surface.

a huge variance in the matching, e.g., Cluster 6 and Cluster 4 show a high percentage of spectral matches, whereas for Clusters 1, 7, and 9 most of the interest points fail the test. This is an interesting find because Cluster 7 represents submerged objects in Parguera 2002 data that are missing in the Parguera 2003 image.

V. CONCLUSION AND DISCUSSION

In this paper, we have described a novel approach for multiscale interest point detection for hyperspectral images. Our algorithm simultaneously adapts to location as well as scale of the point neighborhood. Keeping in mind the size and complexity of hyperspectral images, our development of an interest point operator offers several advantages: It is fast due to the mapping of key computations to the PCA domain, robust due to the projection selection criterion, and allows multiscale detection using the DoG kernel.

Interest points using the extremas of Laplacian have shown several advantages like invariance to geometric transformation, scale, and viewpoint changes [3], [4], [12]. We used an approximation of this concept in the form of DoG kernels for each spectral layer because of the computational advantage. We addressed several issues specific to time-lapse hyperspectral images in an effort to extend concept of interest points robustly and efficiently for these images. The proposed projection selection algorithm provides spatial smoothing for the hyperspectral image. We also developed methods to preprocess the hyperspectral images for bias and variance by histogram matching that compensates for illumination variances in a simple manner. This step led to an increase of more than 100 matched detections in each image. The algorithm was tested on hyperspectral images from a variety of regions like coastal islands (La Parguera), land–water interface (Chesapeake Bay), rocky terrain (the Cuprite Mining District of Nevada), and plain lands (agricultural field images of Kansas and Oklahoma). These experiments not only tested the performance of the interest point operator for geometric and viewpoint changes,

TABLE II
ANALYSIS OF SPECTRAL CLUSTERING OF INTEREST POINTS

CLUSTER ID (FROM DENDROGRAM)	MARKER IN THE IMAGE (SEE FIGURE 7)	CLUSTER SIZE IN IMAGE PARGUERA 2002	MATCHING* POINTS IN PARGUERA 2003	SPECTRAL MATCH	REPRESENTING REGION IN IMAGE	NEAREST MATCHING CLUSTER FROM DENDROGRAM
1	Red dot	127	82 (64.56)	0	Inland valleys	2
2	Green dot	35	30 (85.7%)	21 (70%)	Inland areas	1
3	Yellow dot	269	185 (68.7%)	62(33.51%)	Inland areas	4
4	Red star	57	36 (63.1%)	35 (88.8%)	Inland areas	3
5	Green star	90	80 (88.88%)	45 (56.25%)	Coastal regions in land	8
6	Yellow star	11	8 (72.7%)	8 (100%)	Near coastal areas	10
7	Red plus	758	407 (53.69%)	53 (13.02%)	Submerged objects in the sea	9
8	Yellow plus	27	20 (74.07%)	8 (40%)	Offshore islands	5
9	Green plus	37	28 (74.67%)	2 (7.14%)	Coastal land	7
10	Red cross	1	1	0	-	6

but also tested its robustness in presence of noise and outliers (objects present in one image but not in its pair). The percentage of matching interest points (52% to 75%) is adequately dense for correspondence extraction and feature-based image registration.

The proposed interest point for hyperspectral images differs from conventional monochromatic interest points in few ways. Unlike Lowe's and other algorithms for monochromatic images, the orientation of the hyperspectral interest points was not established in this paper. This is because spatial gradient field is difficult to be estimated in hyperspectral images. In monochromatic images, the spatial gradient field is estimated from finite difference approximation of the neighboring pixels. However, in hyperspectral images, we were dealing with a high-dimensional vector at each pixel location, which resulted in a high-dimensional spatial gradient vector which is not very reliable and useful based on our experience. Therefore, the notion of orientation as understood in the monochromatic interest point literature was not adopted in the hyperspectral case. Moreover, due to the lack of orientation information, we could not define spatial descriptors for hyperspectral interest points as it was described for monochromatic images. Spatial descriptor assigns a feature-vector to a monochromatic interest point by creating a histogram of orientation distribution in its local neighborhood. This gives each interest point a unique identity and helps in its matching and recall. However, we have proposed spectral descriptors for hyperspectral interest points that characterize each interest point based on the spectral information at its location and scale. We have also shown the usefulness of spectral descriptors in matching time lapse images. The spectral signatures are useful to hyperspectral images because they can be computed with few computations, and they can serve as discriminating feature-vector for classifying interest points. Our preliminary attempt at hierarchical clustering has shown some promising result as described in Section IV, where it was shown that the detected points can be refined and categorized to increase the selectivity of matching, and to permit additional image analysis. Future work includes

the development of more discriminant descriptors as well as the use of spatial neighborhood constraints.

APPENDIX

Consider a hyperspectral pixel vector with N spectral bands as $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), h_2(\mathbf{x}), \ldots, h_N(\mathbf{x})]^{\mathrm{T}}$. Then, (1) can be written as

$$
\begin{aligned}
D(\mathbf{x}, \sigma) &= f_l \left(\mathrm{DoG}(\mathbf{x}, \sigma, l) \right) \\
&= f_l \left(\mathbf{h}(\mathbf{x}) \otimes g(k\sigma) - \mathbf{h}(\mathbf{x}) \otimes g(\sigma) \right) \\
&= f_l \left(\begin{bmatrix} h_1(\mathbf{x}) * g(k\sigma) \\ h_2(\mathbf{x}) * g(k\sigma) \\ \vdots \\ h_N(\mathbf{x}) * g(k\sigma) \end{bmatrix} - \begin{bmatrix} h_1(\mathbf{x}) * g(\sigma) \\ h_2(\mathbf{x}) * g(\sigma) \\ \vdots \\ h_N(\mathbf{x}) * g(\sigma) \end{bmatrix} \right)
\end{aligned}
$$
(A1)

where $*$ represents linear convolution operation in space. Consider $\mathbf{h}(\mathbf{x})$ being decomposed into its principal components, therefore convolution operation in (A1) can be written as

$$
\mathbf{h}(\mathbf{x}) \otimes g(\sigma) = \begin{bmatrix} h_1(\mathbf{x}) * g(\sigma) \\ h_2(\mathbf{x}) * g(\sigma) \\ \vdots \\ h_N(\mathbf{x}) * g(\sigma) \end{bmatrix} = \begin{bmatrix} \left(\sum_i \alpha_i(\mathbf{x})\lambda_i v_{1i} \right) * g(\sigma) \\ \left(\sum_i \alpha_i(\mathbf{x})\lambda_i v_{2i} \right) * g(\sigma) \\ \vdots \\ \left(\sum_i \alpha_i(\mathbf{x})\lambda_i v_{Ni} \right) * g(\sigma) \end{bmatrix}
$$
(A2)

where i is the index of the principal components, α_i denotes the projection of the image on ith principal component and λ_i and $\mathbf{v}_i = [v_{1i}, v_{2i}, \ldots, v_{Ni}]^{\mathrm{T}}$ are the eigenvalues and eigenvectors for ith principal component. Note that in (A2), for each spectral band l, λ_i and v_{li} are scalars, and thus by distributive law of convolution $g(\sigma)$ can be convolved with $\alpha_i(\mathbf{x})$ for each principal components. Moreover, therefore in our notation, we can write $\mathbf{h}(\mathbf{x}) \otimes g(\sigma) = \sum_i (\alpha_i(\mathbf{x}) \otimes g(\sigma))\lambda_i \mathbf{v}_i$.

Thus

$$
\begin{aligned}
D(\mathbf{x}, \sigma) &= f_l \left(\mathbf{h}(\mathbf{x}) \otimes g(k\sigma) - \mathbf{h}(\mathbf{x}) \otimes g(\sigma) \right) \\
&= f_l \left(\sum_i \left(\alpha_i(\mathbf{x}) * g(k\sigma) - \alpha_i(\mathbf{x}) * g(\sigma) \right) \lambda_i \mathbf{v}_i \right) \\
&= f_l \left(\sum_i \mathrm{DoG} \left(\alpha_i(\mathbf{x}), \sigma, i \right) \lambda_i \mathbf{v}_i \right).
\end{aligned}
$$

ACKNOWLEDGMENT

The authors would like to thank CenSSIS colleagues J. M. Duarte-Carvajalino at the University of Puerto Rico at Mayaguez for providing data, and to Prof. R. Radke at Rensselaer for offering valuable advice on this paper.

REFERENCES

[1] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*. Reading, MA: Addison-Wesley, 1993.

[2] Y. Bentoutou, N. Taleb, K. Kpalma, and J. Ronsin, "An automatic image registration for applications in remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 2127–2137, Sep. 2005.

[3] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[4] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *Int. J. Comput. Vis.*, vol. 37, no. 2, pp. 151–172, Jun. 2000.

[5] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 12–16, Jan. 2002.

[6] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[7] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17–28, Jan. 2002.

[8] J. S. Pearlman, P. S. Barry, C. C. Segal, J. Shepanski, D. Beiso, and S. L. Carman, "Hyperion, a space-based imaging spectrometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1160–1173, Jun. 2003.

[9] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vis. Conf.*, Sep. 1988, pp. 147–151.

[10] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[11] C. S. Kenney, M. Zuliani, and B. S. Manjunath, "An axiomatic approach to corner detection," in *Proc. IEEE Comput. Soc. Conf. CVPR*, 2005, vol. 1, pp. 191–197.

[12] T. Lindeberg, "Feature detection with Automatic scale selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, Nov. 1998.

[13] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *Int. J. Comput. Vis.*, vol. 60, no. 1, pp. 63–86, Oct. 2004.

[14] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2003, pp. II-257–II-263.

[15] T. Kadir, A. Zisserman, and M. Brady, "An affine invariant salient region detector," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 228–241.

[16] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 1986.

[17] M. Brown and D. Lowe, "Invariant features from interest point groups," in *Proc. Brit. Mach. Vis. Conf.*, Cardiff, U.K., 2002, pp. 656–665.

[18] A. E. Abdel-Hakim and A. A. Farag, "CSIFT: A SIFT descriptor with color invariant characteristics," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2006, vol. 2, pp. 1978–1983.

[19] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky, "Fast geodesic active contours," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1467–1475, Oct. 2001.

[20] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, no. 4, pp. 532–540, Apr. 1983.

[21] J. M. P. Nascimento and J. M. B. Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.

[22] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: Wiley-Interscience, 2000.

[23] J. C. Harsanyi and C. I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, Jul. 1994.

[24] Y. Rubner, C. Tomasi, and L. J. Guibas, "The Earth mover's distance as a metric for image retrieval," *Int. J. Comput. Vis.*, vol. 40, no. 2, pp. 99–121, Nov. 2000.

[25] H. Ling and K. Okada, "Diffusion distance for histogram comparison," in *Proc. Comput. Vis. Pattern Recog.*, 2006, vol. 1, pp. 246–253.

[26] H. Wu, B. Norum, J. Newmark, B. Salzberg, C. M. Warner, C. DiMarzio, and D. Kaeli, "The CenSSIS image database," in *Proc. 15th Int. Conf. Sci. Statistical Database Manage.*, 2003, pp. 117–126.

[27] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv. (CSUR)*, vol. 24, no. 4, pp. 325–376, Dec. 1992.

Amit Mukherjee received the B.E. degree in instrumentation and control engineering from the University of Pune, Pune, India, in 2001 and the M.S. degree in electrical engineering from the University of Houston, Houston, TX, in 2004. He is currently working toward the Ph.D. degree with the Department of Electrical Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY.

His research interest primarily includes signal and image processing, computer vision, pattern recognition, and neural networks.

Miguel Velez-Reyes (S'81–M'92–SM'00) received the B.S. degree in electrical engineering from the University of Puerto Rico at Mayagüez (UPRM), Mayagüez, Puerto Rico, in 1985 and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1988 and 1992, respectively.

Since 1992, he has been a Professor with the Department of Electrical and Computer Engineering, UPRM, where he is also currently the Director of the Laboratory of Applied Remote Sensing and Image Processing. He is the Director of the UPRM Tropical Center for Earth and Space Studies: a NASA University Research Center and the Associate Director of the Center for Subsurface Sensing and Imaging Systems: an NSF Engineering Research Center led by Northeastern University, Boston, MA. He has held faculty internship positions with AT&T Bell Laboratories, the Air Force Research Laboratories, and the NASA Goddard Space Flight Center. He has over 60 publications in journals and conference proceedings. His teaching and research interests include the areas of model-based signal processing, system identification, parameter estimation, and remote sensing using hyperspectral imaging.

Dr. Vélez-Reyes is a member of the Academy of Arts and Sciences of Puerto Rico and of the Tau Beta Pi, Sigma Xi, and Phi Kappa Phi honor societies. In 1997, he was one of 60 recipients from across the U.S. and its territories of the Presidential Early Career Award for Scientists and Engineers from the White House.

Badrinath Roysam (SM'89) received the B.Tech. degree in electronics engineering from the Indian Institute of Technology, Madras, India, in 1984 and the M.S. and D.Sc. degrees from Washington University, St. Louis, MO, in 1987, and 1989, respectively.

Since 1989, he has been with Rensselaer Polytechnic Institute, Troy, NY, where he is currently a Professor in the Department of Electrical Computer and Systems Engineering. He is an Associate Director of the Center for Subsurface Sensing and Imaging Systems (CenSSIS)—a multiuniversity NSF-sponsored engineering research center—and a Codirector of the Rensselaer Center for Open Source Software. He also holds an appointment in the Biomedical Engineering Department. His ongoing projects are in the areas of 2-D, 3-D, and 4-D biomedical image analysis, biotechnology automation, optical instrumentation, high-speed and real-time computing architectures, and parallel algorithms.

Dr. Roysam is a member of the Microscopy Society of America, the International Society for Analytical Cytology, the Society for Neuroscience, and the Association for Research in Vision and Ophthalmology. He is an Associate Editor for the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING and IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE.