Unsupervised Detection of Sub-pixel Objects in Hyper-spectral Images via Diffusion Bases

Alon Schclar[1] and Amir Averbuch[2]

[1]*School of Computer Science, The Academic College of Tel-Aviv Yaffo, POB 8401, Tel Aviv 61083, Israel*

[2]*School of Computer Science, Tel Aviv University, POB 39040, Tel Aviv 69978, Israel*

Keywords: Image Processing, Subpixel, Segmentation, Anomaly Detection, Unsupervised, Sub-pixel Detection, Diffusion Bases, Dimensionality Reduction, Hyper-spectral Sensing.

Abstract: Sub-pixel objects are defined as objects which due to their size and due to the resolution of the camera occupy a fraction of a pixel or partially span adjacent pixels. Unsupervised detection of sub-pixel objects can be highly useful in areas such as medical imaging, and surveillance, to name a few. *Hyper-spectral* images offer extensive intensity information by describing a scene at hundreds and even thousands of wavelengths. This information can be utilized to obtain better sub-pixel detection results compared to those that are obtained using RGB images. Usually, only a small number of wavelengths contain the information that is required for the detection. Furthermore, the intensity images of many wavelengths are noisy and contain very little information. Accordingly, hyper-spectral images must be pre-processed first in order to extract the information that is needed for the sub-pixel detection. This extraction process produces an image where each pixel is represented by a small number of features which allows the application of fast and efficient detection algorithms. In this paper we propose the Diffusion Bases (DB) dimensionality reduction algorithm in order to derive the essential features for the sub-pixel detection. The effectiveness of the DB algorithm facilitates the application of a very simple algorithm for the detection of sub-pixel objects in the feature space. The proposed approach does not assume any distribution of the background pixels. We demonstrate the proposed framework for the detection of cardboard objects in airborne hyper-spectral images of a desert terrain.

1 INTRODUCTION

Sub-pixel objects appear in images due to their relative small size compared to the resolution of the camera and the camera's distance from the objects. This is commonly found in images of areas taken from high-altitude cameras.

For example, in rescue missions of people lost in open areas a camera is mounted on an airplane to record images of the searched area. These images are then analyzed by real-time object detection algorithms. Due to the time-critical nature of such situations, large areas need to be covered quickly which can be achieved by flying a rescue plane at a high altitude. However, the higher the plane flies, the smaller the objects appear in the images which may cause the searched objects to occupy only a fraction of a pixel. Thus, fast algorithms for the detection of sub-pixel objects are required in such scenarios. Hyper-spectral cameras can assist the detection of sub-pixel objects. Such cameras capture an image in the visible spectrum as well as the invisible spectrum i.e. the infra-red and ultra-violet spectrum sub-ranges. Body heat, for example, can only be seen in the infra-red range. Detection algorithms can combine the infra-red information with the visible information to produce fast and accurate detection results.

Unfortunately, processing large scale hyper-spectral images incurs a computational cost that is too high for most applications due to the high number of wavelengths. This is commonly known as the *curse of dimensionality*. Furthermore, hyper-spectral images usually contain noise due to poor lighting conditions and physical conditions at the time the images were taken. Hyper-spectral images also contain redundant information since the number of wavelengths is much higher than the actual degrees of freedom of the data. Consider for example a task that separates green objects from blue objects using an off-the-shelf digital camera. In this case, the camera will produce, in addition to the red and blue channels, a red channel, which is unnecessary for this task. This phenomenon

Schclar, A. and Averbuch, A.
Unsupervised Detection of Sub-pixel Objects in Hyper-spectral Images via Diffusion Bases.
DOI: 10.5220/0008201604960501
In *Proceedings of the 11th International Joint Conference on Computational Intelligence (IJCCI 2019)*, pages 496-501
ISBN: 978-989-758-384-1

is usually unavoidable since general purpose cameras are used. Moreover, constructing a mission-specific camera according to *a-priori* knowledge of the wavelengths that are necessary for the task is unfeasible. Consequently, a hyper-spectral image must initially be pre-processed in order to remove the noise and redundant information and extract the smallest amount of information needed to facilitate the efficient detection of the sought-after objects. This process is commonly known as *feature extraction* or *dimensionality reduction* and its output concisely describes each original data item - in our case pixels - by a small number of attributes instead of original high number of values.

Every substance in nature has a unique *spectral signature* that is described by the substance reflectance values at the different wavelengths of the spectrum. Spectral signatures are often described by hundreds or even thousands of wavelengths depending on the hyper-spectral acquisition instrument. Considering each signature as a vector, the number of wavelengths defines the *dimensionality* of the signature.

Methods for detecting sub-pixel objects in hyper-spectral images can be divided into two categories - supervised and unsupervised. Supervised methods commonly utilize the uniqueness of spectral signatures. Specifically, these methods use *a-priori* spectral information of the sought-after objects. Unsupervised techniques, on the other hand, do not utilize any *a-priori* information and rely on the fact that the spectral signature of sub-pixel objects differs from those of their neighboring pixels. The algorithm proposed in this paper falls into the latter category.

The proposed algorithm initially applies the recently introduced *Diffusion Bases* (DB) dimensionality reduction algorithm (Schclar and Averbuch, 2015; Schclar and Averbuch, 2017b; Schclar and Averbuch, 2017a) to extract a small number of features. The DB algorithm is chosen since it efficiently captures non-linear inter-wavelength correlations and produces a low-dimensional representation in which the amount of noise is drastically reduced. The main contribution of this paper is the application of the DB algorithm since it produces a low-dimensional represtation where sub-pixel objects appear as pixels that are *substantially* different from their neighboring pixels. The sub-pixel objects can then be detected using a very simple procedure.

This paper is organized as follows: in section 2 we present a survey of related work on detection of sub-pixel objects in hyper-spectral images. The DB algorithm is described in section 3. In section 4 we introduce the two phase sub-pixel object detection algorithm. Section 5 contains experimental results and concluding remarks are given in section 6.

2 RELATED WORKS

Subpixel segments are also regarded in the literature as anomalies. Different approaches have been proposed to detect subpixel segments.

The Reed-Xiaoli detector (Reed and Yu, 1990) is considered the baseline to many algorithms that followed. Specifically, this detector assumes that the background follows a Gaussian distribution and uses the Mahalanobis distance of each pixel to its neighbors to detect sub-pixel objects. The covariance matrix is calculated globally for the entire background. In (Zhao et al., 2015) a variation of the Reed-Xiaoli detector is proposed for situations in which the Gaussian distribution assumption does not hold globally. Namely, the covariance matrix is only calculated in a local neighborhood of each pixel. A kernelized version of the Reed-Xiaoli detector is presented in (Kwon and Nasrabadi, 2005). The pixels are implicitly embedded in high-dimensional space where they can be detected using a simple Euclidean distance. This approach generalized the baseline Reed-Xiaoli to cases where straightforward application of the Euclidean distance fails to detect sub-pixel segments.

More recently, Ma *et al.* (Ma et al., 2018) proposed a Deep Belief Network (DBN) to detect sub-pixel objects. An autoencoder is used to extract high-level features. Then, sub-pixels segments are determined according to their weighted distance to their neighboring pixels. Olson *et al.* (Olson et al., 2018) proposed a manifold learning approach coupled with sampling and out-of-sample extension to model the background. Sampling can derive a background model that is more accurate than using the entire image since the sample will be dominated by background pixels.

3 THE DB DIMENSIONALITY REDUCTION ALGORITHM

The *DB* algorithm (Schclar and Averbuch, 2015) utilizes and preserves non-linear inter-coordinate correlations to reduce the dimensionality of a given dataset (it is dual to the *Diffusion Maps* algorithm (Coifman and Lafon, 2006; Schclar, 2008; Schclar et al., 2010)). Since the uniqueness of each signature is also inherent in its inter-wavelength correlations, the DB algorithm is highly effective as a pre-processing tool for hyper-spectral images (Schclar and Averbuch, 2017b). Specifically, applying the DB algorithm to hyper-spectral images reduces their dimensionality while maintaining the vital spectral information of the captured scene.

A hyper-spectral image can be represented as individual monochromatic images - one for each wavelength. Each image is composed of the reflectance values of the scene at a specific wavelength range. The DB algorithm first constructs a graph in which the wavelength images constitute the vertices and the weights are determined according to a fast decay similarity function. Next, it calculates the eigenvectors of the graph Laplacian. The eigenvectors are sorted in descending order according to the magnitude of their eigenvalues and only the eigenvectors whose eigenvalues are above a given threshold are maintained. These eigenvectors capture the *non-linear* inter-wavelength variability of the original data. The selected eigenvectors are used as an orthonormal system on which the pixels of the original image are projected. The projected values constitute the extracted features.

Although baring some similarity to PCA, this process produces better results than PCA due to: (a) its ability to capture non-linear correlations within the data by local exploration of each coordinate; and (b) its robustness to noise. Furthermore, the DB algorithm is more general than PCA and they coincide when the weights in the graph are determined using the inner product weight function.

We denote by $H = \{x_i\}_{i=1}^m, x_i \in \mathbb{R}^n$ the dataset of the pixels in the hyper-spectral image. Let $x_i(j)$ be the j-th coordinate (the reflectance value at the j-th wavelength) of x_i, $1 \le j \le n$. We define the vector $x'_j \triangleq (x_1(j), \ldots, x_m(j))$ as the j-th coordinate of all the points in H i.e. the image corresponding to the j-th wavelength. We denote the set of wavelength images by

$$H' = \{x'_j\}_{j=1}^n. \quad (1)$$

Let $w_\sigma(x_i, x_j)$, be a weight function which measures the similarity between the points x_i and x_j in H'. Ideally, $w_\sigma(x_i, x_j) \to 1$ when x_i and x_j are similar and $w_\sigma(x_i, x_j) \to 0$ otherwise, where σ defines the size of the local neighborhood of each data point. We denote by w_σ the weight matrix that is composed of all pairwise similarities in H. A common choice of w_σ is the Gaussian, however, other kernels that follow a similar decay property can be chosen. A detailed discussion regarding the choice of w_σ and σ is given in (Schclar and Averbuch, 2015).

A Markov transition matrix P is constructed by normalizing the sum of each row in the matrix w_σ to be 1:

$$p(x'_i, x'_j) = \frac{w_\sigma\left(x'_i, x'_j\right)}{d(x'_i)}, \ i, j = 1, \ldots, n$$

where $d(x'_i) = \sum_{j=1}^n w_\sigma\left(x'_i, x'_j\right)$ is the degree of x'_i. Next, the eigen-decomposition of $p\left(x'_i, x'_j\right)$ is calculated

$$p(x'_i, x'_j) \equiv \sum_{k=1}^n \lambda_k r_k(x'_i)\, l_k(x'_j)$$

where $\{l_k\}$ and $\{r_k\}$ denote the left and the right eigenvectors of P, respectively, and $\{\lambda_k\}_{k=1,\ldots,n}$ are the eigenvalues of P in descending order of magnitude. We make use of the eigenvalue decay property of the eigen-decomposition and construct the orthonormal system containing only the first ν eigenvectors $B \triangleq \{r_k\}_{k=1,\ldots,\nu}$. These eigenvectors capture the *non-linear* directions with the highest variability of the coordinates of the original dataset H. We project the original data H onto the orthonormal system B. Let $H_B = \{g_i\}_{i=1}^m, g_i \in \mathbb{R}^\nu$ be the set of these projections where $g_i = (\langle x_i, r_1\rangle, \ldots, \langle x_i, r_\nu\rangle)$, $i = 1, \ldots, m$ and $\langle\cdot,\cdot\rangle$ denotes the inner product operator. H_B is the dimension reduced representation of H and the coordinates of each pixel contain the extracted features. The *DB* algorithm is summarized in Algorithm 1.

4 THE DB SUB-PIXEL OBJECT DETECTION ALGORITHM

We introduce a simple and efficient algorithm for the detection of sub-pixel objects in hyper-spectral images. First, the algorithm normalizes each wavelength image so that its values will be in the range $[0,1]$. This is necessary since different wavelength sensors can produce values at different scales. We denote by $\widetilde{H} = \{\widetilde{x}_i\}_{i=1}^m, \widetilde{x}_i \in \mathbb{R}^n$ the set of normalized wavelength images. Next, features are extracted from $\widetilde{H}$ using the *DB* algorithm be letting $H' = \widetilde{H}$ in Algorithm 1. The extracted features capture the required information for the detection. Let

$$\widetilde{H}_B = \{\widetilde{g}_i\}_{i=1}^m, \widetilde{g}_i \in \mathbb{R}^\nu$$

be the set of pixels whose coordinates are composed of the extracted features. We denote by $\widetilde{H_B}^k = \left(\widetilde{g}_i^k\right)_{i=1,\ldots,m}$ the values of the k-th feature of all the pixels. We normalize each $\widetilde{g}_i^k$ to be in $[0,1]$ and denote the the result by $\widehat{H_B}^k = \left(\widehat{g}_i^k\right)_{i=1,\ldots,m}$. We refer to $\widehat{H_B}^k$ as the k-th *feature image.*

Algorithm 1: The Diffusion Basis algorithm.

Input:
H' **-** A dataset where each wavelength image is a data item
w_σ **-** A similarity function
σ **-** The local neighborhod size of each point
ν - The number of extracted features
Output:
H_B - The reduced dimension representation
======================
DiffusionBasis(H', w_σ, σ, ν)

1. Calculate the weight function $w_\sigma\left(x'_i, x'_j\right)$, $i, j = 1, \ldots n$ where x'_i and x'_j are the i-th and j-th wavelength images.
2. Construct a Markov transition matrix P by normalizing each row in w_σ to sum to 1:
$$p\left(x'_i, x'_j\right) = \frac{w_\sigma\left(x'_i, x'_j\right)}{d\left(x'_i\right)}$$
where $d\left(x'_i\right) = \sum_{j=1}^{n} w_\sigma\left(x'_i, x'_j\right)$.
3. Perform eigen-decomposition of $p\left(x'_i, x'_j\right)$
$$p\left(x'_i, x'_j\right) \equiv \sum_{k=1}^{n} \lambda_k r_k\left(x'_i\right) l_k\left(x'_j\right)$$
where the left and the right eigenvectors of P are given by $\{l_k\}$ and $\{r_k\}$, respectively, and $\{\lambda_k\}$ are the eigenvalues of P in descending order of magnitude.
4. Project the original data H onto the orthonormal system $B \triangleq \{r_k\}_{k=1,\ldots,\nu}$ to obtain
$$H_B = \{g_i\}_{i=1}^{m}, g_i \in \mathbb{R}^\nu$$
where
$$g_i = \left(\langle x_i, r_1\rangle, \ldots, \langle x_i, r_\nu\rangle\right), i = 1, \ldots, m, r_k \in B, 1 \le k \le \nu$$
and $\langle\cdot,\cdot\rangle$ is the inner product.
5. **return** H_B.

The effectiveness of the DB feature extraction allows us to employ a very simple algorithm in order to detect sub-pixel objects. The algorithm makes use of the fact the

sub-pixel objects are substantially different from the hyper-pixels in their local neighborhood. This difference is apparent in a number of their features and it is visible when inspecting some of the individual feature images. This is due to the difference in their correlations with the vectors in the diffusion basis which stems from the difference between the spectral signature of the sub-pixel object and the spectral signature of its neighboring hyper-pixels.

The detection of sub-pixel objects is composed of the following steps. We define the α-neighborhood of $\widehat{g}_i^k$ to be
$$\alpha\left(\widehat{g}_i^k\right) \triangleq \left\{\widehat{g}_j^k \,|\, \|i - j\| \le \alpha\right\}$$
where $\|i - j\|$ denotes the distance between pixel i and pixel j. Next, we compute the number of pixels in $\alpha\left(\widehat{g}_i^k\right)$ whose differences from $\widehat{g}_i^k$ are above a given threshold τ_1. We denote this number by
$$\Delta_\alpha\left(\widehat{g}_i^k\right) \triangleq \left|\left\{\widehat{g}_j^k : \left|\widehat{g}_i^k - \widehat{g}_j^k\right| > \tau_1, \, \widehat{g}_j^k \in \alpha\left(\widehat{g}_i^k\right)\right\}\right|.$$
If the size of $\Delta_\alpha\left(\widehat{g}_i^k\right)$ is larger than a given threshold τ_2 then $\widehat{g}_i^k$ is classified as a sub-pixel object *candidate*. τ_2 determines the number of pixels that are required to be different from $\widehat{g}_i^k$ in its neighborhood, in order for $\widehat{g}_i^k$ to be classified a sub-pixel object candidate.

Finally, a pixel i is classified as a *sub-pixel object* if it is a candidate in at least two different feature images. This final requirement prevents the misclassification of noisy pixels as sub-pixel objects.

5 EXPERIMENTAL RESULTS

In order to simulate sub-pixel objects, twenty four cardboards were placed in a desert terrain. The color of the cardboards resembled the color of the sand in order to make their detection harder using only

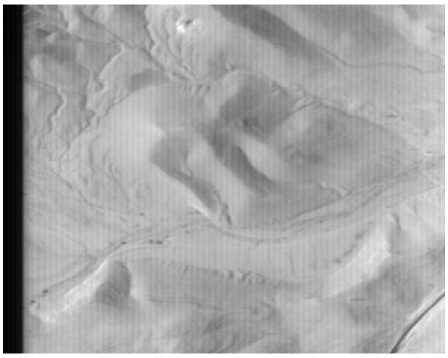

Figure 1: The WAV of a hyper-spectral image of a mountain terrain which contains 24 sub-pixel objects. The sub-pixel objects are visually undetectable.

the visible spectrum. A 121 wavelengths 300×300 hyper-spectral image of the terrain was taken from a high altitude airplane. The altitude was determined according to the camera resolution and the size of the cardboards so that the cardboards will appear as sub-pixel objects in the image.

In order to display the geometry (objects, background, etc.) of the hyper-spectral image we use a 300×300 gray image which is derived as the average of the 121 wavelength images. We refer to it as the *wavelength-averaged-version* (WAV) of the hyper-spectral image.

The similarity function that was used by the DB algorithm was the Gaussian kernel $w_\sigma(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$. where σ was chosen according to the procedure that is described in (Schclar and Averbuch, 2015).

The image and the results are given in Fig. 1-4. Figure 1 shows the WAV of the image. Figures 2 and 3 show the 35^{th} and 50^{th} wavelengths, respectively. The noise in the image is due to atmospheric conditions at the time the image was taken. These problems can be found in most of the wavelengths. The sub-pixel segments are manifested in the image as isolated points. Figure 4 displays the 2^{nd} feature image $\widehat{G}^2$ with squares around the sub-pixel segments that were found. The number of extracted features was empirically set to $\nu = 6$. The sub-pixel detection was obtained using $\tau_1 = 0.04$, $\tau_2 = 3$. These values were found empirically. The algorithm detected *all* twenty four sub-pixel objects.

6 FUTURE RESEARCH

The values of ν, τ_1 and τ_2 were found empirically. Naturally, the optimal values for these parameters are

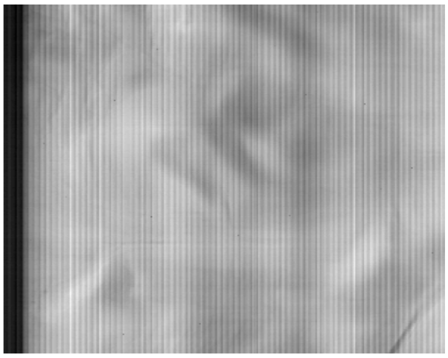

Figure 2: The 35^{th} wavelength of Fig. 1.

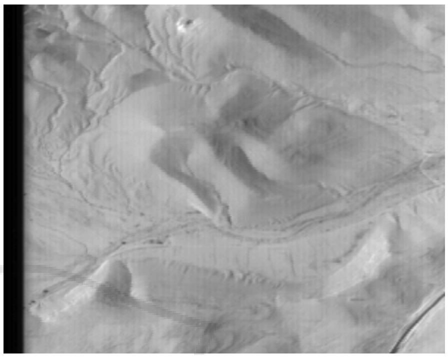

Figure 3: The 50^{th} wavelength of Fig. 1.

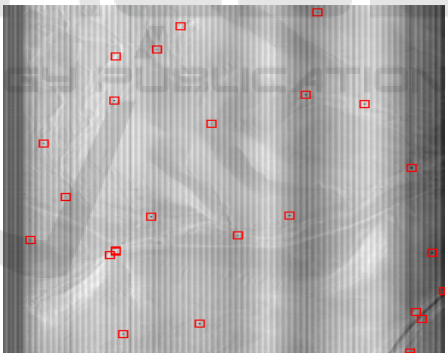

Figure 4: $\widehat{G}^2$ of Fig. 1 with squares around the detected sub-pixel segments. The parameters that were used are: $\nu = 6, \tau_1 = 0.04$, $\tau_2 = 3$.

data driven (similarly to choosing ε in (Schclar and Averbuch, 2015)) i.e. they depends on the given hyper-spectral image .Automatic choice of these parameter can simplify and accelerate the detection of sub-pixel objects. A rigorous way for choosing them is currently being investigated by the authors.

The results in section 5 were obtained using a Gaussian kernel. However, it is shown in (Coifman and Lafon, 2006) that any positive semi-definite kernel may be used for the dimensionality reduction. Rigorous analysis of families of kernels to facilitate the derivation of an optimal kernel for a given image H is currently an open problem that should be investigated.

REFERENCES

Coifman, R. R. and Lafon, S. (2006). Diffusion maps. *Applied and Computational Harmonic Analysis: special issue on Diffusion Maps and Wavelets*, 21:5–30.

Kwon, H. and Nasrabadi, N. M. (2005). Kernel rx-algorithm: A nonlinear anomaly detector for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 43:388–397.

Ma, N., Peng, Y., Wang, S., and Leong, P. H. W. (2018). An unsupervised deep hyperspectral anomaly detector. *Sensors*, 18(3):693.

Olson, C. C., Judd, K. P., and Nichols, J. M. (2018). Manifold learning techniques for unsupervised anomaly detection. *Expert Systems and Applications*, 91(C):374–385.

Reed, I. S. and Yu, X. (1990). Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(10):1760–1770.

Schclar, A. (2008). *A Diffusion Framework for Dimensionality Reduction*, pages 315–325. Springer US, Boston, MA.

Schclar, A. and Averbuch, A. (2015). Diffusion bases dimensionality reduction. In *Proceedings of the 7th International Joint Conference on Computational Intelligence, IJCCI 2015, Lisbon, Portugal, November 12-14, 2015.*, pages 151–156.

Schclar, A. and Averbuch, A. (2017a). A diffusion approach to unsupervised segmentation of hyper-spectral images. In *Computational Intelligence - 9th International Joint Conference, IJCCI 2017 Funchal-Madeira, Portugal, November 1-3, 2017 Revised Selected Papers*, pages 163–178.

Schclar, A. and Averbuch, A. (2017b). Unsupervised segmentation of hyper-spectral images via diffusion bases. In *Proceedings of the 9th International Joint Conference on Computational Intelligence, IJCCI 2017, Funchal, Portugal, 1-3 November, 2017.*, pages 151–156.

Schclar, A., Averbuch, A., Hochman, K., Rabin, N., and Zheludev, V. (2010). A diffusion framework for detection of moving vehicles. *Digital Signal Processing,*, 20(1):111–122.

Zhao, C., Wang, Y., Qi, B., and Wang, J. (2015). Global and local real-time anomaly detectors for hyperspectral remote sensing imagery. *Remote Sensing*, 7(4):3966–3985.