REGRESSION FRAMEWORK FOR BACKGROUND ESTIMATION IN REMOTE SENSING IMAGERY

James Theiler and Brendt Wohlberg

Los Alamos National Laboratory, Los Alamos, NM 87545 USA

ABSTRACT

A key component in any target or anomaly detection algorithm is the characterization of the background. We investigate several approaches for estimating the background level at a given pixel, based on both the local neighborhood around that pixel and on the global context of the full image. By framing this as a regression problem, we can compare a variety of background estimation schemes, from standard signal processing approaches long used in the hyperspectral image analysis community to more sophisticated nonlinear approaches that have recently been developed in the image processing community. These comparisons are performed on a range of images including single band, standard red-green-blue, eight-band WorldView-2, and 126-band hyperspectral HyMap imagery.

Index Terms— Target, Anomaly, Background, Regression, Hyperspectral, Image, Processing

1. INTRODUCTION

To detect small targets in large images, it is important that the background be well estimated. A number of standard detection algorithms – adaptive matched filter (AMF) [1, 2, 3], adaptive coherence estimator (ACE) [4, 5], finite target matched filter (FTMF) [6] – begin with the assumption that the background is a multivariate Gaussian, whose mean $\boldsymbol{\mu}$ and covariance R are estimated from the full dataset. This assumption is highly restrictive, and amounts to estimating the background at a given pixel by the global mean of all the pixels in the entire image. A natural way to improve this estimate is to compute the mean (and, possibly, the covariance) from a moving window centered on the pixel of interest. The venerable RX algorithm [7, 8] takes this approach.

Matteoli *et al.* [9] recommend a smaller window for estimating the mean $\boldsymbol{\mu}$ and a larger window for estimating the covariance R. The limit of large covariance window size – so that a global covariance estimator is used along with the local estimates of mean – is sometimes called "semilocal" estimation. Cohen *et al.* [10, 11] investigated a variety of algorithms and a variety of background estimators. Even taking into account the contaminating effects of the target on the background, including both pixel phasing and optical blurring, it was found that using smaller annuli produced better detectors.

But all of these methods are based on local averaging to estimate $\boldsymbol{\mu}$. Our hypothesis is that better performance may be achieved by treating the problem as a (linear or nonlinear) regression. For a given pixel at position i, we want to estimate its value (which is vector-valued for multispectral and hyperspectral imagery) based on the pixels in its neighborhood $\mathcal{N}_i$ (this is usually an annular region that does not include the pixel itself, and in some cases avoids a "guard ring" of pixels in the immediate neighborhood of i). The traditional approach is to estimate $\mathbf{x}_i$ with the average of those neighbors:

$$\hat{\mathbf{x}}_i = \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} \mathbf{x}_j \tag{1}$$

In our regression approach, we seek a function f, learned from the entire image, such that

$$\hat{\mathbf{x}}_i = f(\{\mathbf{x}_j\}_{j \in \mathcal{N}_i}) \; . \tag{2}$$

Background estimation – though it's typically not called that – is also an important part of classical image processing. It arises both in noise reduction [12] (where the "background" at a pixel location is the noise-free pixel value) and in the "inpainting" problem [13, 14, 15] (where one seeks to fill in – or to "paint in" – pixel values that are for some reason unavailable). In both cases, one seeks to exploit spatial structure in the image in order to estimate a pixel value based on the values of its neighbors. Our problem is perhaps a little closer to the inpainting problem. For each pixel in the scene, we remove the pixel, in-paint a replacement, and measure the difference between the original pixel and its inpainted replacement. If the two are very different, then that pixel is a candidate for a target location.

2. REGRESSION ALGORITHMS

The simplest algorithm, and the baseline for the studies presented here, is the local averaging algorithm in Eq. (1) that estimates the value of the central pixel as the arithmetic average of the pixels in the annulus.

Our nonlinear algorithms are also based on an annulus around a pixel, but they exploit statistics that we can learn from the rest of the image regarding the relationship of the of pixels in an annulus to the pixel in the center. The simplest of these nonlinear algorithms is based on k-nearest neighbors. The idea is to identify the k patches in the image that are most similar to the annulus that surrounds the pixel of interest. The center pixels for those k patches are then prototypes for the pixel of interest. In our implementation, we took $k = 3$, and used the median of those prototypes as our estimator for the pixel of interest. (We also tried using a mean, and obtained essentially identical results.)

We also considered two other nonlinear algorithms that employed more sophisticated combinations of these nearest neighbor patches. In the non-local mean (NLM) [16] algorithm, the simple mean (or median) is replaced by a *weighted* mean, with the weights chosen to depend on the dis-similarity of the neighbor patches with the patch of interest. In particular, if d_{im} is the dis-similarity of patches $\mathcal{N}_i$ and $\mathcal{N}_m$ (measured as the mean of squared differences of the pixels in those annular patches), then the weight is given by

Both authors are funded by the Laboratory Directed Research and Development (LDRD) program at Los Alamos National Laboratory.

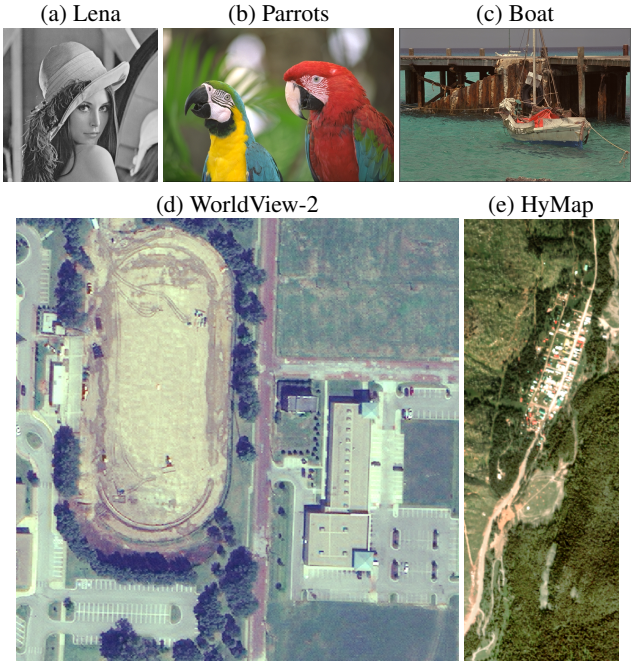

Fig. 1. Three of the images (a,b,c) used in this study are popular choices for comparing standard image processing algorithms, and two of the images (d,e) are remote sensing images. (a) Lena is a single band (black and white) image, and (b) Parrots and (c) Boat are both three-band (red, green, blue) images. (d) WorldView-2 is a multispectral image with eight bands, and (e) HyMap is a hyperspectral data cube, with 126 spectral channels.

$w_{im} = \exp(-\alpha d_{im})$, and the estimator for the pixel of interest becomes

$$\hat{\mathbf{x}}_i = \sum_m w_{im}\mathbf{x}_m \Big/ \sum_m w_{im}. \tag{3}$$

We have found the the performance does not depend too sensitively on α, but it is nonetheless a parameter that must be estimated. We do this by taking a smaller sub-image and searching for an α that optimizes $||\hat{\mathbf{x}}_i - \mathbf{x}_i||^2$, averaged over the sub-image.

Finally, we consider a non-local sparse (NLS) representation. This is similar to NLM, but for NLS weights are fit directly so that $\mathcal{N}_i \approx \sum_m w_{im}\mathcal{N}_m$, subject to a sparsity constraint on w_{im} [17]. That is, the annular patch of interest is fit with a linear combination of near neighbor patches, and that linear combination is then applied to the central pixels in order to provide an estimate $\hat{\mathbf{x}}_i$.

All of these non-local algorithms require a search for near neighbors in patch space; this is much more computationally expensive than the purely local averaging. We were able to speed up the computation by using the Generalized PatchMatch algorithm of Barnes *et al.* [18].

Although we ultimately care about target or anomaly detection performance, our focus here will be on the intermediate question of how precisely we can estimate the background; specifically, we will use the mean squared error (MSE) of the background estimate relative to the actual background value. This is the most straightforward from a regression point of view. Suppose $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ is the error in an estimate $\hat{\mathbf{x}}$ of $\mathbf{x}$. Here, $\mathbf{e}$ is a vector with a number of elements equal to the number of spectral channels in the image. If we write $\mathbf{e}_k$ as the kth pixel of n total, then

$$\text{MSE} = \frac{1}{n}\sum_{k=1}^{N} \mathbf{e}_k^T\mathbf{e}_k = \langle \mathbf{e}^T\mathbf{e}\rangle. \tag{4}$$

Often this is expressed in terms of decibels with respect to the variance of the image:

$$\text{SNR(dB)} = -10\log_{10}\langle \mathbf{e}^T\mathbf{e}\rangle + 10\log_{10}\langle \mathbf{x}^T\mathbf{x}\rangle. \tag{5}$$

3. IMAGERY

We used five images in this study; three of them, shown in Fig. 1(a,b,c), are popular choices for standard image processing algorithms. Two of them, shown in Fig. 1(d,e) are obtained from remote sensing imagers; one is eight-band multispectral, and one is 126-band hyperspectral.

Lena is a single-band ("black and white") 512×512 pixel image (with an interesting history [19]) that includes edges, textures, gradients. The Parrots and Boat images are three band (red, green, blue) color 512×768 pixel images [20]. A WorldView2 [21, 22] image was provided to us by Digital Globe; this 512×512 crop is from a much larger image taken of Omaha, Nebraska. We used the blind reflectance dataset from the RIT blind test [23] as our hyperspectral sample; there are 126 channels and 280×800 pixels.

4. EXPERIMENTS WITH SINGLE BAND IMAGES

For a single band image (or for a single band in a multispectral image), all the information is spatial, and by comparing local averaging to nonlocal methods, we test how much information is in the spatial structure of an image. Intuitively, we know that images have a lot of spatial structure, and our eye is well practiced at picking out subtle spatial clues. But is remote sensing imagery qualitatively different from other kinds of imagery – *e.g.*, the images that are taken of people or their pets and that show up in places like ImageNet [24]? We know that remote sensing routinely employs spectral ranges that are unavailable in snapshot photography, but the other half of the question is: are the spatial statistics of remote sensing imagery fundamentally different from those of ImageNet style images? It is not obvious that the kinds of images for which image processing tools such as nonlocal mean were designed will have utility in remote sensing images. The single band experiments test this question. What we observe in Table 1 is that the nonlocal estimators consistently outperform local averaging.

This table also shows that a smaller annular patch generally outperforms larger patches; this result was also observed by Cohen *et al.* [10] in their experiments. The effect is quite striking for local averaging, whereas the nonlocal estimators were seen to be more robust to increasing patch size. For the remaining experiments, we use a small annulus: $\mathcal{N}_i$ is a 3×3 patch of pixels with the center pixel missing.

5. MULTISPECTRAL AND HYPERSPECTRAL IMAGERY

As we move from single band to RGB, it is feasible to treat each pixel as a three-component vector of values, and to apply the nonlocal regression to these vectors. The results, shown in Table 2, echo those seen in Table 1 for the single band images.

To avoid searching in higher dimensional spaces, an alternative approach is to treat each of the bands independently, and to apply

Table 1. Average SNR(dB) for various regression schemes applied to single bands of the images; shown here are results are based on two different patch sizes.

(a) 3×3 annular patch

	Avg.	kNN	NLMean	NLSparse
Lena (band 2)	17.47	22.78	23.10	20.88
Parrots (band 2)	18.93	25.39	24.76	24.13
Boat (band 2)	12.42	18.46	19.06	17.24
WV-2 (band 3)	12.34	16.67	16.93	15.40
HyMap (band 20)	17.79	21.39	22.23	23.44

(b) 5×5 annular patch

	Avg.	kNN	NLMean	NLSparse
Lena (band 2)	14.29	20.14	19.68	19.28
Parrots (band 2)	15.73	22.60	21.33	23.18
Boat (band 2)	10.09	15.73	15.16	15.88
WV-2 (band 3)	9.37	14.38	13.94	14.76
HyMap (band 20)	12.39	18.05	18.95	23.81

(c) 7×7 annular patch

	Avg.	kNN	NLMean	NLSparse
Lena (band 2)	12.46	19.18	18.33	18.98
Parrots (band 2)	14.20	21.54	19.41	22.98
Boat (band 2)	9.24	14.73	13.44	15.15
WV-2 (band 3)	8.23	13.42	11.87	14.99
HyMap (band 20)	9.81	16.77	16.06	23.68

the estimation one component at a time. This is a simpler approach (though more expensive), and as Table 2(b) shows, its performance is comparable to (perhaps even better than) the approach in Table 2(a) based on treating all three bands together. Further, since we have already demonstrated that individual spectral components (*i.e.*, individual bands) are better predicted using nonlinear methods, we can be confident that this result will obtain for multispectral and hyperspectral images as well.

The fact that single band performance improvements can always be applied band-by-band to a full image provides a path to achieving lower MSE on hyperspectral images. The nonlinear methods, however, are more expensive, and this expense becomes significant in an image with 100+ spectral bands.

We have employed a hybrid algorithm in which PCA is applied first to the image, and nonlinear band-by-band estimation of the first few components is performed. Simple local averaging is performed on the remaining bands, the results are combined, and the PCA is inverted. Table 3 shows that the nonlinear improvements to only a few principal components are enough to substantially improve the performance vis-a-vis local averaging for all the bands. Table 3 also compares band-by-band regression for the three principal components to a scheme that treats those three bands as a three-component vector. As we saw in the RGB case, the band-by-band approach gives better performance. Here, the bands are nominally independent (they are uncorrelated, at least) and so we might expect band-by-band to work particularly well.

6. FUTURE WORK

Although mean squared error provides one way to evaluate performance, we ultimately prefer an approach that more directly maps to target detection performance. Future efforts will consider the utility of these nonlocal estimators in the context of specific target detection scenarios. Target implantation [25, 10, 11] provides a direct approach for producing reliable scenarios, when the target signature is known, but volume-based approaches have also been suggested [26, 27], particularly for anomaly detection (when the target signature is unknown).

Also, the work described here concentrated on estimating the mean value of the estimated background. While it is possible to employ a locally varying mean with a global covariance, there is opportunity for further improvement by considering methods to locally estimate the covariance as well [28, 29, 30, 31].

An alternative approach to local patches for nonlinear regression is to use kernel methods [32, 33, 34]; it is not clear how these methods compare, both in terms of theoretical similarities and in terms of practical target detection performance.

Table 2. Average SNR(dB) for various regression schemes applied to three bands of the images; results here are for a 3×3 annular patch.

(a) bands together

	Avg.	kNN	NLMean	NLSparse
Parrots	20.13	25.60	25.47	24.66
Boat	12.05	17.62	18.63	16.49
WV-2 (bands 1,2,3)	12.94	16.06	17.54	15.64
HyMap (bands 1,2,3)	16.04	19.49	20.24	21.83

(b) band-by-band

	Avg.	kNN	NLMean	NLSparse
Parrots	20.13	26.48	25.26	23.59
Boat	12.05	18.10	17.93	16.50
WV-2 (bands 1,2,3)	12.94	17.24	17.79	15.95
HyMap (bands 1,2,3)	16.04	19.66	19.13	22.45

Table 3. Average SNR(dB) for local averaging and nonlocal k-Nearest-Neighbors applied to all the bands in the images, but using a hybrid scheme that applies the nonlocal method only to the first three principal components, and applies local averaging to the rest.

	Avg.	Hybrid kNN band-by-band	Hybrid kNN bands together
Parrots	20.13	26.53	25.82
Boat	12.05	18.13	17.83
WV-2	16.80	20.80	19.75
HyMap	19.63	22.63	21.28

7. REFERENCES

[1] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerospace and Electronic Systems*, vol. 10, pp. 853–863, 1974.

[2] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerospace and Electronic Systems*, vol. 22, pp. 115–127, 1986.

[3] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerospace and Electronic Systems*, vol. 28, pp. 208–216, 1992.

[4] L. L. Scharf and L. T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," *Proc. 30th Asilomar Conference on Signals, Systems, and Computers*, pp. 1114–1117, 1996.

[5] S. Kraut, L. L. Scharf, and R. W. Butler, "The Adaptive Coherence Estimator: a uniformly most-powerful-invariant adaptive detection statistic," *IEEE Trans. Signal Processing*, vol. 53, pp. 427–438, 2005.

[6] A. Schaum and A. Stocker, "Spectrally selective target detection," *Proc. ISSSR (International Symposium on Spectral Sensing Research)*, p. 23, 1997.

[7] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1760–1770, 1990.

[8] A. D. Stocker, I. S. Reed, and X. Yu, "Multi-dimensional signal processing for electro-optical target detection," *Proc. SPIE*, vol. 1305, pp. 218–231, 1990.

[9] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE A&E Systems Magazine*, vol. 25, pp. 5–27, 2010.

[10] Y. Cohen, D. G. Blumberg, and S. R. Rotman, "Subpixel hyperspectral target detection using local spectral and spatial information," *J. Applied Remote Sensing*, vol. 6, no. 1, pp. 063508, 2012.

[11] Y. Cohen, Y. August, D. G. Blumberg, and S. R. Rotman, "Evaluating sub-pixel target detection algorithms in hyperspectral imagery," *J. Electrical and Computer Engineering*, vol. 2012, 2012.

[12] P. Milanfar, "A tour of modern image filtering: New insights and methods, both practical and theoretical," *IEEE Signal Processing Magazine*, vol. 30, pp. 106–128, Jan 2013.

[13] M. Bertalmío, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 2000, pp. 417–424, ACM Press/Addison-Wesley Publishing Co.

[14] Z. Xu and J. Sun, "Image inpainting by patch propagation using patch sparsity," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1153–1165, 2010.

[15] B. Wohlberg, "Inpainting by joint optimization of linear combinations of exemplars," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 75–78, Jan. 2011.

[16] A. Buades, B. Coll, and J.-M. Morel, "Image denoising by non-local averaging," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, March 2005, vol. 2, pp. 25–28.

[17] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, November 1993, pp. 40–44.

[18] C. Barnes, E. Shechtman, D. B. Goldman, and Adam Finkelstein, "The generalized PatchMatch correspondence algorithm," in *European Conference on Computer Vision (ECCV)*, 2010.

[19] D. C. Munson, "A note on Lena," *IEEE Trans. Image Processing*, vol. 5, 1996.

[20] "Kodak lossless true color image suite," `http://r0k.us/graphics/kodak/`.

[21] G. Marchisio, F. Pacifici, and C. Padwick, "On the relative predictive value of the new spectral bands in the WorldView-2 sensor," *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2723–2726, 2010.

[22] N. T. Anderson and G. B. Marchisio, "Worldview-2 and the evolution of the DigitalGlobe remote sensing satellite constellation: introductory paper for the special session on WorldView-2," *Proc. SPIE*, vol. 8390, 2012.

[23] D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive, and S. Hager, "Development of a web-based application to evaluate target finding algorithms," *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2, pp. 915–918, 2008.

[24] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2009.

[25] W. F. Basener, E. Nance, and J. Kerekes, "The target implant method for predicting target difficulty and detector performance in hyperspectral imagery," *Proc. SPIE*, vol. 8048, pp. 80481H, 2011.

[26] J. Theiler and D. Hush, "Statistics for characterizing data on the periphery," *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 4764–4767, 2010.

[27] J. Theiler, "Ellipsoid-simplex hybrid for hyperspectral anomaly detection," *Proc. 3rd IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2011.

[28] P. V. Villeneuve, H. A. Fry, J. Theiler, B. W. Smith, and A. D. Stocker, "Improved matched-filter detection techniques," *Proc. SPIE*, vol. 3753, pp. 278–285, 1999.

[29] C. E. Caefer, J. Silverman, O. Orthal, D. Antonelli, Y. Sharoni, and S. R. Rotman, "Improved covariance matrices for point target detection in hyperspectral data," *Optical Engineering*, vol. 47, pp. 076402, 2008.

[30] G. Cao and C. A. Bouman, "Covariance estimation for high dimensional data vectors using the sparse matrix transform," in *Advances in Neural Information Processing*, Daphne Koller, Yoshua Bengio, Dale Schuurmans, Léon Bottou, and Aron Culotta, Eds. 2009, pp. 225–232, MIT Press.

[31] L. Bachega, J. Theiler, and C. A. Bouman, "Evaluating and improving local hyperspectral anomaly detectors," *Proc. 40th IEEE Applied Imagery Pattern Recognition (AIPR) Workshop*, 2011.

[32] D. Tax and R. Duin, "Data domain description by support vectors," in *Proc. ESANN99*, M Verleysen, Ed., Brussels, 1999, pp. 251–256, D. Facto Press.

[33] H. Kwon and N. M. Nasrabadi, "Kernel-based subpixel target detection in hyperspectral images," *Proc. IEEE Intl. Conf. Neural Networks*, pp. 717–721, 2004.

[34] H. Kwon and N. M. Nasrabadi, "Kernel RX: a new nonlinear anomaly detector," *Proc. SPIE*, vol. 5806, pp. 35–46, 2005.