Efficient Clustering-based Noise Covariance Estimation for Maximum Noise Fraction

Soumyajit Gupta
Dept. of Computer Science
University of Texas, Austin
Email: smjtgupta@utexas.edu

Chandrajit Bajaj
Dept. of Computer Science
University of Texas, Austin
Email: bajaj@cs.utexas.edu

***Abstract*—Most hyperspectral images (HSI) have important spectral features in specific combination of wave numbers or channels. Noise in these specific channels or bands can easily overwhelm these relevant spectral features. Maximum Noise Fraction (MNF) by Green *et al.* [1] has been extensively studied for noise removal in HSI data. The MNF transform maximizes the Signal to Noise Ratio (SNR) in feature space, thereby explicitly requiring an estimation of the HSI noise . We present two simple and efficient Noise Covariance Matrix (NCM) estimation methods as required for the MNF transform. Our NCM estimations improve the performance of HSI classification, even when ground objects are mixed. Both techniques rely on a superpixel based clustering of HSI data in the spatial domain. The novelty of our NCM's comes from their reduced sensitivity to HSI noise distributions and interference patterns. Experiments with both simulated and real HSI data show that our methods significantly outperforms the NCM estimation in the classical MNF transform, as well as against more recent state of the art NCM estimation methods. We quantify this improvement in terms of HSI classification accuracy and superior recovery of spectral features.**

***Keywords*—Hyperspectral image, Maximum noise fraction, Noise covariance estimation, Superpixel, Classification**

I. Introduction

The emergence and development of hyperspectral (HS) remote sensing technology has made it possible to acquire data with large amounts of spatial and spectral information for image analysis applications such as classification, unmixing, subpixel mapping, and target detection [2]. HS imaging is applicable to a wide variety of fields, including agriculture, environment, mineral mapping, surveillance, and chemical imaging of biological tissue [3] [4]. The acquired hyper spectral images (HSI) are often disturbed by radiometric noise such as sensor noise, photon (or shot) noise, calibration error, atmospheric scattering and absorption, which not only degrades the visual quality but also the final HSI analysis or interpretation via image classification techiques. The various aforementioned HSI noise is generally represented as an additive normally distributed (Gaussian), zero-mean random process [5].

It was shown by Green *et al.* [1] that the variance of hyperspectral images did not necessarily reflect the real SNR, due to unequal noise variances in different channels or bands, with noise variance dominating the signal variance in some bands. As a result, a band with small variance does not necessarily mean poor image quality. It may have a high SNR compared to other bands with large variances but low SNR's. In order to deal with this problem, Green *et al.* [1] developed the Maximum Noise Fraction (MNF) transform based on maximization of SNR, so that the transformed principal components are ranked by SNR rather than variance as used in PCA.

One of the major disadvantages of this approach is that the Noise Covariance Matrix (NCM) must be estimated completely and accurately from the data apriori. This is generally difficult to do so, due to the stochastic nature of noise, and as pointed out in [1] [6]. Several papers have been done to address this issue by considering the neighboring spatial information [7] [8], as well as jointly considering neighborhood spatial and spectral information [9] [10], leading to high complexity algorithms just for the NCM estimation.

Our main contributions are two-fold. First, we present a simple method to generate a panchromatic image from the HSI. Second, we propose two variants of NCM estimation approaches (neighborhood and cluster) based on the fast superpixel [11] SLIC algorithm. They depend only on the spatial information of the generated image, and can effectively recover a good estimate of the NCM, achieving good classification scores after MNF denoising. Being spatial based, they have lower computation cost (proportional to the number of pixels in the HSI), compared to spectral based methods. Our results, as demonstrated below, significantly outperforms both the spatial and spectral based approaches in terms of classification accuracy and recovery of spectral features. Both our algorithms remain unaffected with the scaling of the data in terms of size, hence provides a better NCM estimation for large datasets.

II. Theory

A. Notation

We are given a HSI $Y_{orig} \in \mathbb{R}^{W \times H \times S}$, where W, H are the spatial dimensions (width, height) and S is the number of spectral channels. It can be restructured/vectorized into a 2D matrix $Y \in \mathbb{R}^{N \times S}$, where $N = (W \times H)$ is the number of pixels. Each column $y_i, \forall i = [1 : S]$ represents the reshaped image for the i-th spectral band and each row $y_j, \forall j = [1 : N]$ represents the spectral signature for the j-th pixel.

B. Related Work

The main difference between conventional PCA and MNF is that: MNF has a prior step of noise whitening, which needs a good estimate of the NCM. The original MNF method mainly adopts the spatial feature of image to estimate Σ_δ, such as minimum/maximum autocorrelation factor (MAF) by

Switzer and Green [7], causal simultaneous autoregressive and quadratic surface by Nielsen [8].

As shown in studies by Roger [12], Greco *et al*. [13] and Liu *et al*. [9], spatial-based noise estimation method is data-selective and unstable. When HSI has low spatial resolution, the difference between pixels may mainly contain signal. Sometimes, noise with regular (interference) pattern may be considered as signal when spatial features are used in NCM estimation. When the structure/distribution of the noise is unknown, such methods always fall short in estimating Σ_δ.

In HSI there exists correlation between bands. Therefore, high correlations between bands can also be used for noise estimation, as in Liu *et al*. [9] by using uses both spatial and spectral domain. The signal value of current pixel at a band is estimated from the value of adjacent bands of the same pixel and adjacent pixels in the same band, through multiple linear regression. Then difference between the estimated signal value and the raw value of the current pixel, is considered as noise.

The work of Optimized MNF by Gao *et al*. [10] extends the idea of Liu *et al*. [9] by computing noise over small non-overlapping sub-blocks locally, thereby further reducing the influence of spatial features. They solve multiple linear analysis over each sub-block to compute the residual noise, followed by computing the local standard deviation (LSD) for each of those sub-blocks, and binning those values over the range of computed LSD.

C. Superpixel

A superpixel is defined as a group of pixels having similar characteristics. We consider the Simple Linear Iterative Clustering (SLIC) superpixel of Achanta *et al*. [11]. It generates superpixels by clustering pixels based on their color similarity and proximity in the image plane. We use the SLICO version, which adaptively chooses the compactness parameter for each superpixel differently, generating regular shaped groups in both textured and non textured regions alike, and can be used on grayscale and color images. The cost is $O(N)$, hence computationally efficient and real time.

III. Effect of Mixing of Pixels on NCM

Green *et al*. [1] uses the Maximum Autocorrelation Factor (MAF) [7] transform to obtain Σ_δ:

$$\Sigma_\delta = \frac{1}{2} Cov(Y_{j,i} - Y_{j+\Delta,i}) \tag{1}$$

where Δ denotes spatial shift in pixel. In Eq.1, the subtractions between neighboring pixels is performed to estimate Σ_δ. If the spatial distribution of ground objects keeps getting complex as shown in Fig. 1, NCM estimates are very bad, hence classification results after MNF transform will be seriously affected as shown in Fig. 2.

IV. Generating Panchromatic Image for SLIC

A HSI contains large number of continuous spectral bands with narrow bandwidth. To extract spatial features in a fast and simple way, we use a synthesized panchromatic image inspired by the work in Zhang *et al*. [14]. This generated panchromatic image $I \in \mathbb{R}^{W \times H}$ is used as input to the SLIC algorithm.

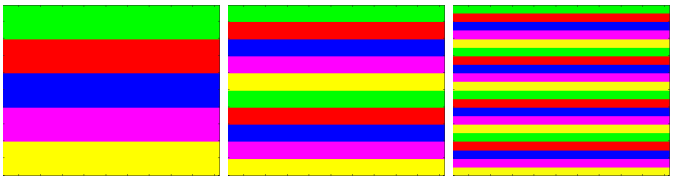

Fig. 1: Ground truth image of five materials of synthetic HSI over three levels of spatial rearrangement of pixels.

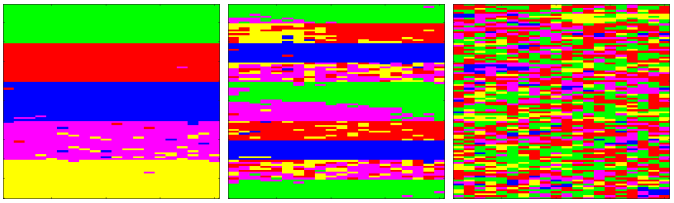

Fig. 2: Result of classification of the sample HSI after MNF denoising. K-Means is used for assigning labels. As classes keep on mixing, the performance of MNF degrades severely due to poor estimation of noise, hence poor NCM estimate.

$$I = w_r I_r + w_g I_g + w_b I_b \tag{2}$$

where I_r, I_g, I_b are the spectral bands of the HSI with band centers corresponding to the red, green, and blue bands. In our experiments, the weights w_r, w_g, w_b are set to $0.06, 0.63, 0.27$ as per values suggested in [14], which are perceptually optimized on human visual data. For HSI measured in the InfraRed (IR) region, there are no perceptually feasible red, green or blue bands. In such cases, we use the image I of the sample corresponding to the band used for normalization.

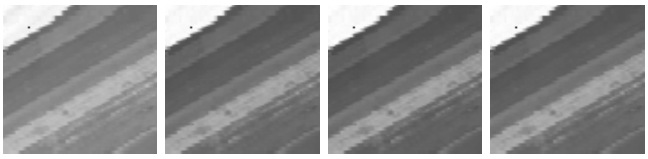

Fig. 3: Left to Right: Original images corresponding to red (665.59 nm), green (589.31 nm), and blue (491.90 nm) bands and Simulated Panchromatic image for the Salinas-A dataset.

V. Methods: Improved NCM Estimators

Spatial based noise estimators are fast but inaccurate when the data gets mixed. This issue is circumvented by considering spectral based noise estimators at the expense of increased computation cost. Another issue hampering the performance of spatial estimators like in Eq. 1, is the spatial dimensions of the data. For large vectorized images, there is absolutely no correlation between pixels at row skips even if they belong to the same material. Our spatial-based estimators handles both cases of accurately estimating noise at low computation cost, as well as being unaffected by spatial dimension. It computes the noise estimates from the 3D cube Y_{orig}, and then vectorizes it to Y, thereby maintaining proper correlations between neighboring pixels in the spatial domain.

A. Neighborhood Based

This noise estimator of Alg. 1 is inspired by earlier spatial based methods, where instead of just looking to the right neighbor, all surrounding 8 neighbors of a pixel are accounted for computing the noise level for that pixel. To avoid estimating noise from neighbors of different material, only the ones that share the same superpixel label are considered. The noise is then computed as the difference between the current pixel's value and the weighted sum of values of its neighbors with same superpixel label.

Algorithm 1 NCM-A($I \in \mathbb{R}^{W\times H}, Y_{orig} \in \mathbb{R}^{W\times H\times S}$)

$\bar{I} = SLIC(I)$
for each pixel $(i,j) \in \bar{I}$ **do**
  for each neighbor (i',j') of pixel (i,j) **do**
    $ct = 0$; $elem = \Phi$
    if $label(\bar{I}(i,j)) == label(\bar{I}(i',j'))$ **then**
      $ct = ct + 1$
      $elem[ct] = Y_{orig}(i',j',:)$
    $\delta_{orig}(i,j,:) = Y_{orig}(i,j,:) - \frac{1}{ct}(\sum_{k=1}^{ct} elem[k])$
Output: $\delta_{orig} \in \mathbb{R}^{W\times H\times S}$ reshaped into $\delta \in \mathbb{R}^{N\times S}$

B. Cluster Based

This noise estimator of Alg. 2 is inspired by the classical K-means, where cluster centers are the true representation of the class spectrum, and all data within that class are some deviations away from the center due to noise and other effects. Once the superpixel labels are assigned, the mean spectrum of each label k is calculated as $\mu(k)$. The noise is then computed as the difference between the current pixel's value and the mean spectrum of the same label.

Algorithm 2 NCM-B($I \in \mathbb{R}^{W\times H}, Y_{orig} \in \mathbb{R}^{W\times H\times S}$)

$\bar{I} = SLIC(I)$
for each label $k \in K$ **do**
  $\mu(k)$ = mean of pixels with label k)
for each pixel $(i,j) \in \bar{I}$ **do**
  $k = label(i',j')$
  $\delta_{orig}(i,j,:) = Y_{orig}(i,j,:) - \mu(k)$
Output: $\delta_{orig} \in \mathbb{R}^{W\times H\times S}$ reshaped into $\delta \in \mathbb{R}^{N\times S}$

VI. Experiments

In this section, we provide simulations to showcase the effectiveness of our twin algorithms. We analyze results on both synthetic images created from real world spectrum as well as on public dataset, to show the robustness in presence of noise from any unknown distribution.

A. Data Description

For the synthetic image shown in Fig. 1, the spectrum contains reflectance values collected from the AVIRIS scene over Moffett Field [1], CA in 1997. We chose regions from five classes namely Light vegetation, Rock, Dense vegetation, Dry soil and Wet soil. We chose this dataset because it is real-world and noise is included in the spectrum, hence no bias due to prior knowledge of the noise distribution.

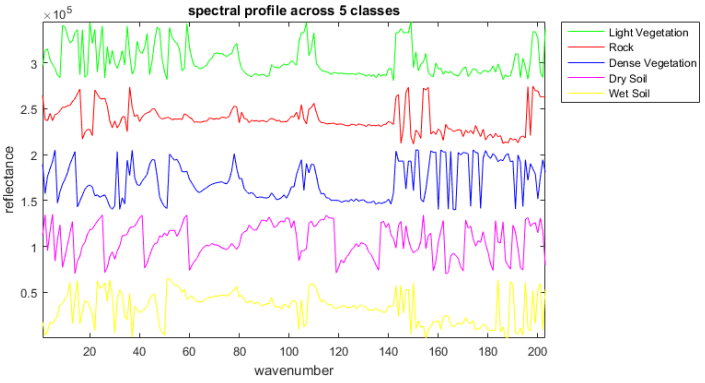

Fig. 4: Clean Spectrum of five classes from the Salinas-A dataset. Light and Dense Vegetation contains similarity across their spectrum, same goes for Dry and Wet Soil.

For experimentation, the data is arranged in stratified layers from each class in chunks of 20×20 pixels, with each spectrum containing 203 channels and 5 classes, resulting in an HSI of size $100 \times 20 \times 203$ of reflectance values.

B. Classification Metrics

We chose PCA for dimension reduction followed by K-means clustering to be the classification algorithm. Of course other complex algorithms for pure pixel classification or end-member extraction can be employed, but those are out of scope for this work. The number of components for PCA was chosen as the number of bands that accounted for 97.5% of the total SNR after MNF transform. Loss is measured as the number of misclassified pixels after MNF denoising. We see that the

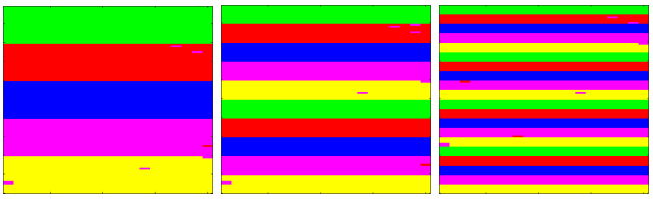

Fig. 5: Result of classification of the sample HSI after MNF denoising with NCM-A Alg. 1. Across levels, the classification accuracy of a pixel does not degrade much due to proper estimation of noise from neighbors of same class.

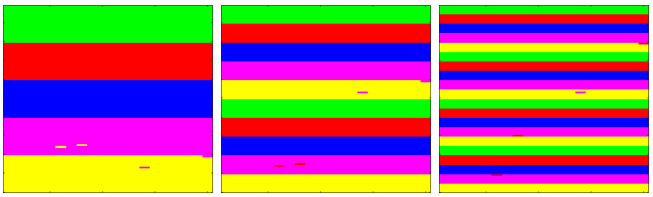

Fig. 6: Result of classification of the sample HSI after MNF denoising with NCM-B Alg. 2. Across levels, the classification accuracy of a pixel does not degrade much due to proper estimation of noise from the mean spectrum of its superpixel.

[1]Data available at http://aviris.jpl.nasa.gov/html/aviris.freedata.html

number of misclassified samples in Fig. 5, 6 is drastically less compared to the MNF with naive noise approximation in Fig. 2. Even though the number of intermixing of layers increased, both our algorithms were able to accurately approximate the noise, hence leading to better HSI denoising.

C. Different Variations of Mixing

In order to further study the effect of mixing of objects, we shuffled the data blocks around to introduce more class skips along both directions, with patterns which occur in real-world datasets like Indian Pines or Salinas. We see that in these cases too, our approach is able to handle the spatial mixing of boundaries accurately. The two class boundary results are shown in Fig. 7, 8. Similarly, the multiple class boundary results are shown in Fig. 9, 10 respectively.

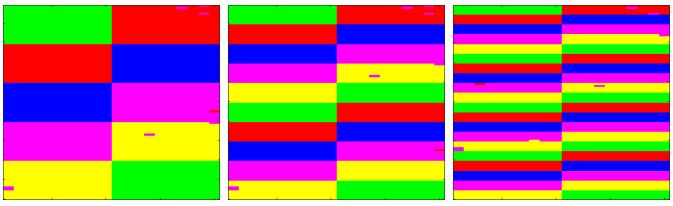

Fig. 7: Result of classification of the sample HSI after MNF denoising with NCM-A Alg. 1

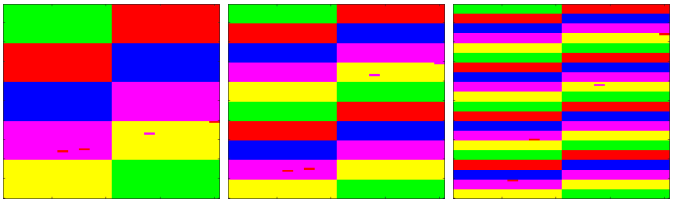

Fig. 8: Result of classification of the sample HSI after MNF denoising with NCM-B Alg. 2

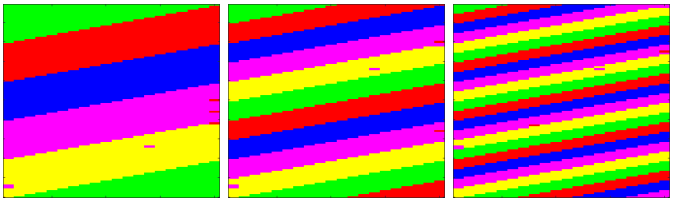

Fig. 9: Result of classification of the sample HSI after MNF denoising with NCM-A Alg. 1

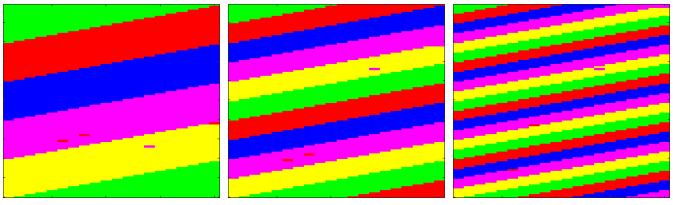

Fig. 10: Result of classification of the sample HSI after MNF denoising with NCM-B Alg. 2

D. Comparison to Spectral based approaches

Here we show the benefit of our approaches compared to the spectral approaches, which are fast but time consuming. We ran the same experiment using Optimized MNF of Goa *et al*. [10]. Firstly, solving the system of regression for each pixel of each band takes huge memory and time. Secondly, depending on the spectrum, this method at times becomes numerically unstable, which results in degraded noise estimation. This effect is clearly illustrated in Fig. 11, where at the first level the classification is fair. However, at subsequent levels, due to mixing of objects the regression solver suffers from being badly scaled and close to singular. This happens due to pixels from multiple materials participating in determining the noise level of a pixel from different class.

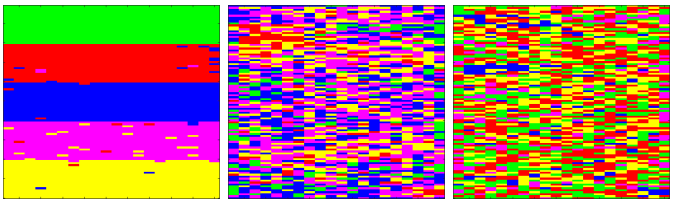

Fig. 11: Result of classification of the HSI after OMNF denoising. With further mixing the results degrade due to instability on solving the regressions.

E. Experiment on Public Dataset

We chose the Salinas-A dataset for experimentation. It was collected by the 224-band AVIRIS sensor over Salinas Valley, CA. It has 86×83 pixels and six classes of radiance spectrum. The 20 water absorption bands were discarded. Fig. 3 shows the generated image for SLIC. The spectrum from six classes are added with random Gaussian noise, which are correlated across bands, but independent of the data. Both the proposed algorithms can accurately recover the true spectral features of the data after MNF denoising, illustrated with low error between true and recovered spectrum in Fig. 12.

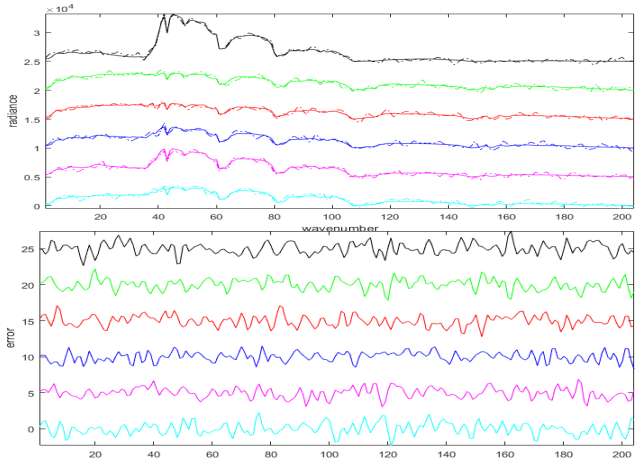

Fig. 12: Salinas-A HSI. Top: True spectrum from six classes in solid, and generated noisy (noise correlated across bands) spectrum in dashed. Bottom: Difference between true and estimated spectrum. Low errors indicate the proposed algorithms accurately estimate the noise introduced in the HSI.

F. Complexity

The cost for computing SLICO superpixel of the image is $O(N)$. NCM-A computes the noise for each pixel by looking into its eight adjacent neighbors, hence takes $O(N)$. NCM-B computes the mean spectra of each superpixel in $O(P)$ time [2], the noise computation for all pixels is again $O(N)$, hence the total time is $O(NP)$. For any HSI data $N \gg P$, hence P is relative constant depending on the dataset used. Hence, the time complexity of the both the algorithms is effectively $O(N)$.

VII. Conclusion

In this work we introduced two spatial based noise estimators for MNF, based on superpixel segmentation of a generated panchromatic image of a HSI. They perform with higher accuracy than the spectral based estimators while having the same computation cost as the spatial based estimators, and with no numerical instability issues. This is relevant in many research areas where the HSI datasets are large, hence a quick and accurate noise estimation method is required to maintain the integrity of MNF denoising.

[2] P is the number of pixels grouped in a superpixel.

References

[1] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on geoscience and remote sensing*, vol. 26, no. 1, pp. 65–74, 1988.

[2] C.-I. Chang, *Hyperspectral imaging: techniques for spectral detection and classification*. Springer Science & Business Media, 2003, vol. 1.

[3] P. Shippert, "Why use hyperspectral imagery?" *Photogrammetric engineering and remote sensing*, vol. 70, no. 4, pp. 377–396, 2004.

[4] Resonon. (2017, may) Hyperspectral imaging applications. [Online]. Available: https://www.resonon.com/applications_main.html

[5] B. Corner, R. Narayanan, and S. Reichenbach, "Noise estimation in remote sensing imagery using data masking," *International Journal of Remote Sensing*, vol. 24, no. 4, pp. 689–702, 2003.

[6] J. B. Lee, A. S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 3, pp. 295–304, 1990.

[7] P. Switzer and A. A. Green, "Min/max autocorrelation factors for multivariate spatial imagery," *Computer science and statistics*, pp. 13–16, 1984.

[8] A. A. Nielsen, "Analysis of regularly and irregularly sampled spatial, multivariate, and multi-temporal data," *Science*, vol. 21, no. 4, pp. 555–567, 1994.

[9] X. Liu, B. Zhang, L. Gao, and D. Chen, "A maximum noise fraction transform with improved noise estimation for hyperspectral images," *Science in China Series F: Information Sciences*, vol. 52, no. 9, pp. 1578–1587, 2009.

[10] L. Gao, B. Zhang, X. Sun, S. Li, Q. Du, and C. Wu, "Optimized maximum noise fraction for dimensionality reduction of chinese hj-1a hyperspectral data," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, p. 65, 2013.

[11] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[12] R. Roger, "Principal components transform with simple, automatic noise adjustment," *International journal of remote sensing*, vol. 17, no. 14, pp. 2719–2727, 1996.

[13] M. Greco, M. Diani, and G. Corsini, "Analysis of the classification accuracy of a new mnf based feature extraction algorithm," in *Remote Sensing*. International Society for Optics and Photonics, 2006, pp. 63 650V–63 650V.

[14] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2579–2591, 2015.