1-7-2004

Performance Analysis of Improved Methodology for Incorporation of Spatial/Spectral Variability in Synthetic Hyperspectral Imagery

Neil W. Scanlan
Rochester Institute of Technology

John Schott
Rochester Institute of Technology

Scott D. Brown
Rochester Institute of Technology

Recommended Citation
Neil W. Scanlan, John R. Schott, Scott D. Brown, "Performance analysis of improved methodology for incorporation of spatial/spectral variability in synthetic hyperspectral imagery", Proc. SPIE 5159, Imaging Spectrometry IX, (7 January 2004); doi: 10.1117/12.509858; https://doi.org/10.1117/12.509858

Performance analysis of improved methodology for incorporation of spatial/spectral variability in synthetic hyperspectral imagery

Neil W. Scanlan, John R. Schott, and Scott D. Brown
Rochester Institute of Technology, Center for Imaging Science
Digital Imaging and Remote Sensing Laboratory
54 Lomb Memorial Drive, Rochester, NY, USA 14623

ABSTRACT

Synthetic imagery has traditionally been used to support sensor design by enabling design engineers to pre-evaluate image products during the design and development stages. Increasingly exploitation analysts are looking to synthetic imagery as a way to develop and test exploitation algorithms before image data are available from new sensors. Even when sensors are available, synthetic imagery can significantly aid in algorithm development by providing a wide range of "ground truthed" images with varying illumination, atmospheric, viewing and scene conditions. One limitation of synthetic data is that the background variability is often too bland. It does not exhibit the spatial and spectral variability present in real data. In this work, four fundamentally different texture modeling algorithms will first be implemented as necessary into the Digital Imaging and Remote Sensing Image Generation (DIRSIG) model environment. Two of the models to be tested are variants of a statistical Z-Score selection model, while the remaining two involve a texture synthesis and a spectral end-member fractional abundance map approach, respectively. A detailed comparative performance analysis of each model will then be carried out on several texturally significant regions of the resultant synthetic hyperspectral imagery. The quantitative assessment of each model will utilize a set of three performance metrics that have been derived from spatial Gray Level Co-Occurrence Matrix (GLCM) analysis, hyperspectral Signal-to-Clutter Ratio (SCR) measures, and a new concept termed the Spectral Co-Occurrence Matrix (SCM) metric which permits the simultaneous measurement of spatial and spectral texture. Previous research efforts on the validation and performance analysis of texture characterization models have been largely qualitative in nature based on conducting visual inspections of synthetic textures in order to judge the degree of similarity to the original sample texture imagery. The quantitative measures used in this study will in combination attempt to determine which texture characterization models best capture the correct statistical and radiometric attributes of the corresponding real image textures in both the spatial and spectral domains. The motivation for this work is to refine our understanding of the complexities of texture phenomena so that an optimal texture characterization model that can accurately account for these complexities can be eventually implemented into a synthetic image generation (SIG) model. Further, conclusions will be drawn regarding which of the candidate texture models are able to achieve realistic levels of spatial and spectral clutter, thereby permitting more effective and robust testing of hyperspectral algorithms in synthetic imagery.

Keywords: Hyperspectral image simulation, DIRSIG, co-occurrence matrix, texture synthesis, image texture

1. INTRODUCTION

1.1 The DIRSIG Model

The Digital Imaging and Remote Sensing Image Generation (DIRSIG) Model is an integrated complex collection of independent first principles based radiation propagation submodels which work in conjunction to produce radiance field images with high radiometric fidelity in the 0.3 – 20 micron region of the electromagnetic spectrum.[1] It is comprised of five main submodels which are designed to allow for a high degree of flexibility and interchangeability of sensor configurations, scene geometry, bi-directional reflectance distribution function (BRDF) predictions, and atmospheric conditions. The DIRSIG model has demonstrated the capability to produce radiometrically rigorous imagery exhibiting properties observed in real imagery such as the simulation of mixed pixels, complex in-situ illumination loadings, and overall spectral statistics[2], however it has been determined that in general, synthetic imagery required significant improvements in order to realistically capture the spatial and spectral complexity present in real image data.[3]

Imaging Spectrometry IX, edited by Sylvia S. Shen, Paul E. Lewis,
Proceedings of SPIE Vol. 5159 (SPIE, Bellingham, WA, 2003)
0277-786X/03/$15 · doi: 10.1117/12.509858

1.2 Motivation

The necessity of characterizing textures correctly in a mathematical and statistical sense, both spatially and spectrally cannot be overstated. One of the most powerful advantages offered through the use of synthetic imagery is the ability to test spatial and spectral exploitation algorithms with great flexibility, while avoiding the logistical and technical challenges associated with field collection efforts. But in order to reap the benefits of synthetic imagery for algorithm testing and development, realistic levels of spatial and spectral clutter must be achieved so that we can reliably estimate how these algorithms will perform on real-world imagery. Otherwise, unrealistically benign backgrounds lacking in spatial and spectral variability and structure will inevitably produce overly optimistic indications of hyperspectral algorithm performance. One must keep in mind that the purpose of mathematically modeling texture (or any phenomenon for that matter) is not to simplify the problem, but rather to be able to understand and include within the model the very complexities that make texture such a challenging problem.

This paper will first briefly describe the theory behind the Single Bandpass (SBP) and Multiple Bandpass (MBP) Z-Score Selection texture models, the Texture Synthesis (TS) model, and the Fraction Map (FM) texture characterization model, after which the texture model performance evaluation metrics that will be used in the assessment of each of these models will be introduced. The results of rendered synthetic imagery of Hyperspectral Digital Imagery Collection Experiment (HYDICE) imagery using the four texture characterization models will then be presented, followed by the quantitative comparative performance analysis of the fidelity of spatial and spectral texture achieved by each model through the application of the GLCM, SCR, and SCM texture metrics. The scope of this paper is limited to the reflective region of the electromagnetic spectrum.

2. METHODOLOGY

2.1 Texture Model Background Theory

Image texture is an intuitive concept with an elusive formal definition that depends on the context of the situation. While the computer graphics community tends to define texture in terms of structural primitives subject to syntactic grammars and placement rules[4], the field of remote sensing describes texture as the structure of the variation in brightness within or between objects of interest. These textures often arise from variations in target reflectance, since most targets are composed of heterogeneous mixtures of materials exhibiting variable spectral reflectance properties. Further affecting the appearance of image texture is topographic effects such as sun-target angles and BRDF and shadowing effects.[5] Due to the diverse interpretations of image texture phenomena, there are accordingly numerous exceptionally complex and widely varying approaches to quantitatively modeling texture, four of which have been adopted for this work and are described below.

2.1.1 SBP Texture Model

The SBP texture model utilizes a z-score selection algorithm within a two-tiered approach to apply spatial-spectral variability to each pixel of the output synthetic image.[2,5] The first tier consists of the generation of a material class map of the image to be rendered in the DIRSIG environment. For this work, the Gaussian Maximum Likelihood (GML) classification algorithm provided the best separability between the eight end member materials of the HYDICE ARM scene. A lookup table (LUT) then assigns a second-tier texture map to each region in the material map in order to introduce spatial variations for a specified wavelength region. The statistical z-score is computed for each pixel in the single bandpass (SBP) texture image in order to drive the selection of a spectral reflectance curve from a large database of ground truth measurements for each material class. The z-scores of the texture image and of each reflectance curve are compared within the texture image bandpass in order to apply spectral variability to each pixel for all wavelength regions of the output SIG image. The SBP texture model clearly requires the existence of thorough and accurate ground truth data in order to adequately characterize the true spatial and spectral variability of all materials present in the scene to be rendered. Although the z-score selection mechanism intelligently relates the spatial-spectral variations of the texture image to those of the spectral database, it only does so for a single bandpass region and thus selects pixel spectra based on the spectral behavior within the specified bandpass region. This carries the potential of selecting spectral curves that are not truly representative of the spectral character of materials in non-correlated spectral bands.

2.1.2 MBP Texture Model

In order to alleviate this concern, the MBP texture model[6] permits the user to select multiple (and ideally uncorrelated) bandpasses so that the spectral character of the pixel can be represented with more fidelity in all spectral bands. This obviously possesses a theoretical advantage over the SBP model, since it will be less likely that a reflectance curve will be chosen that exhibits significantly different behavior than that in the bandpasses from which the curve has been selected. Figure 1 illustrates how a composite weighted z-score is calculated using texture images in three bandpasses. Ideally, the additional bandpasses used should be selected such that they are not well-correlated with the original single bandpass used in the SBP model. Note that the use of additional bandpasses accordingly places the burden on the user to have access to multiple bands of imagery to model the spatial texture of the material class of interest.

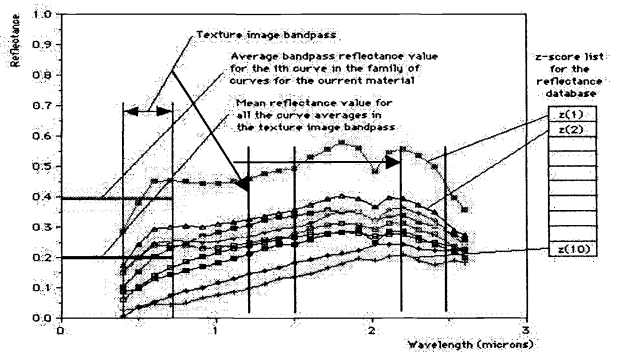


Figure 1: The MBP Texture Model uses multiple texture image bandpasses in order to calculate a composite weighted z-score for each curve (column list, right) and compares these scores to each of the texture image pixel composite z-scores. The curve with the z-score value closest to that of each pixel in the texture image is selected to characterize the spectral behavior of that pixel for the entire spectral extent of the image.

2.1.3 Texture Synthesis Model

The third model to be tested is derived from a texture synthesis-by-analysis technique that iteratively enforces a set of statistical constraints over the output of a complex analytic filter bank by extracting multiresolution scale and space information from a sample monochrome texture.[4] This model has been refined to synthesize multi/hyperspectral textures in the DIRSIG environment.[7] The algorithm begins with a white noise image, and iteratively coerces it to converge into the form of the desired output synthesized texture through a series of statistical constraints, including that of spectral covariance. This method proceeds in an iterative, coarse-to-fine fashion over a variant of the steerable pyramid resampling technique which uses a compact multiresolution representation to analyze the sample texture, by using a set of four oriented complex analytic filters at each level of the pyramid so that local phase information can be used to detect the polarity of edges and boundary transitions. A synopsis of this model is illustrated in Figure 2. This approach has the attractive feature of forcing a solution that matches the desired spectral covariance and spatial correlation statistics in one band. However it cannot assure that the areal spatial patterns within a particular texture region will be reproduced.

2.1.4 Fraction Map Texture Model

The fourth and final model to be tested in this work is derived from the production of fractional abundance maps for end members within the scene to be rendered through the use of a choice of spectral unmixing tools. DIRSIG is able to accept any number of fraction planes and re-mix them in accordance with their fractional abundances on a pixel-by-pixel basis. Each pixel in the resultant DIRSIG image is simply a linear combination of weighted end member spectra, which thereby creates spatial and spectral structure and variability as illustrated in Figure 3.[8] Eight end member fraction maps were produced using simple Linear Spectral Unmixing (LSU) for the HYDICE ARM scene to be rendered. The DIRSIG imagery results using each of the four texture characterization models are presented in Section 3.

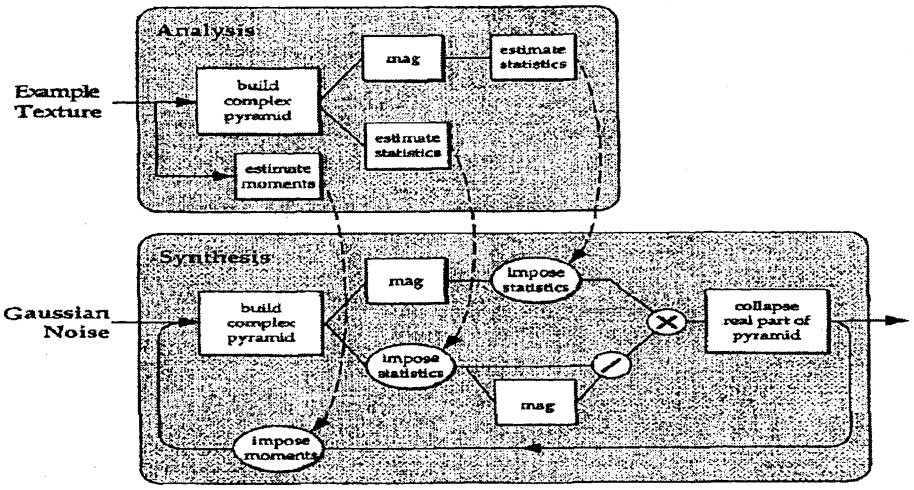


Figure 2: Order of statistical constraint enforcement used within the Texture Synthesis model.

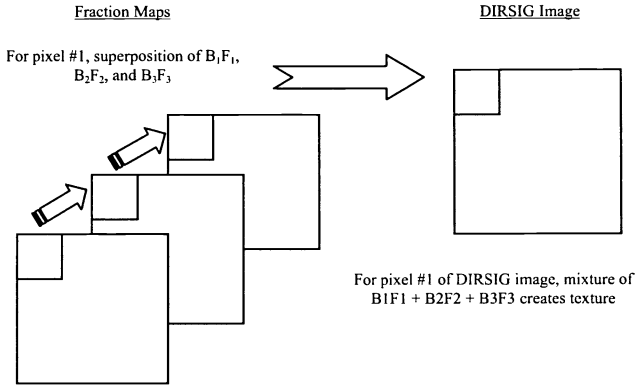


Figure 3: FM Texture Characterization Model. On a pixel-by-pixel basis, the fraction maps corresponding to each end member are re-mixed in order to create spatial and spectral variability in the resultant DIRSIG image.

2.2 Texture Model Performance Metrics

2.2.1 Spatial GLCM Metric

In order to assess the spatial fidelity of each of the SIG images, a metric derived from the popular GLCM measure of spatial texture has been adopted.[9] The GLCM tool has demonstrated a high level of success in the classification, segmentation, and texture feature extraction literature, and has proven to possess among the best discriminatory powers of spatial textural patterns. The GLCM is a parametric description of the spatial arrangement of pixel values presented in a complex matrix form which is essentially a probability density function of all possible co-occurring pixel values within the processing window. The main parameter in its computation is the orientation at which the GLCM is to be calculated. The conjecture is that, if the GLCM-derived statistics are able to discriminate between textural features such that the best results are obtained for most remote sensing applications, then the same method must be sufficient to measure the similarity of texture features between real and synthetic textures in a quantifiable manner. Based on past studies assessing the most effective combinations of GLCM-derived Haralick features and orientation parameters for remotely sensed imagery [10, 11] it was found that the use of the Contrast and Correlation features with a diagonal displacement vector ($(\Delta x, \Delta y) = (1,1)$) provided the most complete description of each texture test region to be measured. The

GLCM, SCR, and SCM metrics have been applied to fifteen texture test regions consisting of homogeneous textures exhibiting within-material class variability and transition textures between material class regions ranging from 10x10 to 35x35 pixels in dimension from the corresponding real and synthetic imagery for each texture model result (see Fig. 6(a)). For the GLCM metric, the Contrast and Correlation texture images for each region were directly compared through the calculation of the absolute difference image and an acceptable variance-threholded difference image. The variance threshold image was defined by performing repeated measurements in the vicinity of each texture test region. Any disparity between the real and DIRSIG texture signatures would therefore be evident in the difference images, and would be indicative of the location and the magnitude of deviation of spatial correlation within each region. This metric offers the advantage of a compact representation of spatial texture signatures for both homogeneous and transition region textures of real and synthetic imagery in order to assess the spatial fidelity of the output DIRSIG imagery on a band-by-band basis. A subset of eight representative spectral bands was selected for the application of all three performance metrics in this work.

2.2.2 Spectral SCR Metric

Researchers in the field of target and anomaly detection in hyperspectral imagery (HSI) commonly employ a measure of the SCR as a threshold for reliable detection of signal patterns in Gaussian clutter.[12] This same measure has been used as an assessment of the similarity of spectral clutter content and complexity within the selected counterpart real and synthetic texture regions in this research. The SCR metric will provide the means to determine if the overall clutter statistics are correct in the synthetic texture in relation to the real image for a given texture type sample. The advantage of this comparison is that the SCR metric generates a single numerical value that will serve as a preliminary measure of how well the overall spectral structure is characterized in the synthetic rendition of the image, before delving into much more detailed spectral analysis involved with the SCM metric. Lastly, many HSI algorithms exploit the measure of SCR in several contexts, so it is essential that this metric is correct for synthesized texture regions if DIRSIG is to support the testing and development of these algorithms.
In its traditional form, the SCR is defined as follows:

$$SCR = \left[\mathbf{b}^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{b}\right]^{1/2} \tag{1}$$

where **M** is the spectral interference (background plus noise) covariance matrix, and **b** is the spectral signature of the "target", which in this case will be the central pixel of the region being examined. In the case of L spectral bands, **b** is a column vector of dimension (1xL), while **M** is given by:

$$\mathbf{M} = \sum_{\mathbf{n=1}}^{\mathbf{N}} \mathbf{x}(n)\mathbf{x}^{\mathbf{T}}(n) = \mathbf{X}^{\mathbf{T}}\mathbf{X} \tag{2}$$

where the matrix $\mathbf{X}^{\mathbf{T}}$ represents the set of N *de-meaned* pixels in the image window under study, and therefore has dimension (LxN), i.e.,

$$\mathbf{X}^{\mathbf{T}} = \left[\mathbf{x}(1) \quad \mathbf{x}(2) \quad \mathbf{x}(3)...\mathbf{x}(N)\right] \tag{3}$$

since each entry **x**(n) is itself a column vector representing the de-meaned spectral signature for a given pixel n. This metric has been applied to each of the fifteen texture regions and for all four texture models. In order to determine the acceptable variance for each of the texture test regions to be studied, the SCR metric has been applied to the same repeated samples as used for the GLCM metric.

2.2.3 Spatial-spectral SCM Metric

The concept of a Generalized Spectral Co-Occurrence Matrix (GSCM) representation has been proposed by Hauta-Kasari et. al.[13] in order to improve texture segmentation results for multispectral imagery. This algorithm generates a co-occurrence matrix that describes the spatial dependency of a quantized spectral domain. Since the GSCM method offers no guarantee that the Self-Organized Mapping (SOM) method of quantization would be carried out in the same manner in the real and synthetic corresponding imagery due to ordering concerns, it has not been adopted for use in this work. However, this concept of using both spatial and spectral information simultaneously was the motivation for the SCM metric that has been used as the third synthetic texture fidelity measure for the HYDICE ARM rendered imagery. The SCM is a novel, simpler yet equally valuable approach to spatial-spectral texture description that has never been used before in the literature on co-occurrence matrices for classification and feature extraction models. The SCM follows a completely analogous process as the GLCM between two *user-specified* spectral bands. The result is a matrix containing

cross-band spatial and spectral co-occurrence information. This tool has been incorporated into the ENVI processing environment such that the user can choose base and comparison spectral bands, which are used as the base and shift windows shown in Figure 4. The parameters of direction vector orientation, Haralick co-occurrence texture features and processing window size are also still available to the user for the SCM and the same parameters as used for the GLCM metric were employed for the application of the SCM.

Base (band A) and shift (band B) windows:

4	3	5
3	5	6
6	4	3

3	5	6
5	6	3
4	3	6

Spectral Co-Occurrence Matrix:

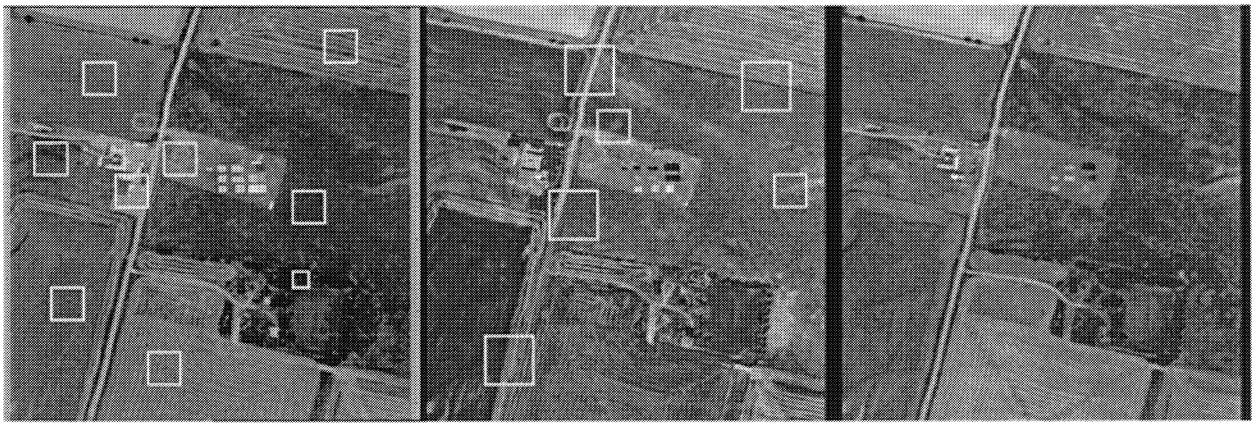
GL	3	4	5	6
3	0	0	2	**1**
4	2	0	0	**0**
5	0	0	0	**2**
6	**1**	**1**	**0**	**0**

Figure 4: Sample ENVI SCM Computation with parameters $\Delta x = 1, \Delta y = 0$.

For the HYDICE ARM data, four band pairs were chosen on which the SCM metric would be performed, with 95%, 5%, -10%, and -40% spectral correlation values. This choice of band pairs was used in order to determine if the spectral correlation in the DIRSIG imagery was preserved in relation to the corresponding real imagery. This sampling of band pairs is considered to be sufficiently representative since it encompasses well-correlated, ill-correlated, negatively ill-correlated, and negatively "well"-correlated spectral structure. An identical testing methodology was used for the SCM as for the GLCM metric, such as the same texture test regions, and the absolute and variance-thresholded difference images were investigated for the Contrast, and Correlation features. The spatial GLCM, spectral SCR, and spatial-spectral SCM metrics will be used in combination to assess the relative performance of the SBP, MBP, TS, and FM texture models in section 4.

3. SCENE SIMULATION RESULTS

The following subsections will present the DIRSIG imagery resulting from the incorporation of the four texture characterization models to be tested in this work. For the sake of brevity, a sampling of only three spectral bands of each image cube will be shown. The corresponding real HYDICE Atmospheric Radiation Measurement (ARM) imagery was originally acquired over the ARM calibration site in Lamont, OK on June 24, 1997. The data was collected using the airborne HYDICE imaging spectrometer flown at an altitude of 3.475 km, which has spectral coverage between 400 – 2,500 nm in 210 spectral bands and GSD of 1.7375 m. A 320x320 pixel subset of Run 29 of the image data was used for the analysis in this paper. Figure 5(a) shows samples of the real HYDICE ARM imagery for comparison purposes with the SIG data. As discussed earlier, the SBP and MBP models rely heavily on accurate and thorough ground truth data in order to generate realistic levels of spatial and spectral clutter. Since we do not intend to test the quality of ground truth data measurements in this work, it was decided to create image-derived "ground truth" reflectance spectra to represent the best-case scenario in which hundreds of reflectance spectra can be produced to characterize a given material class.



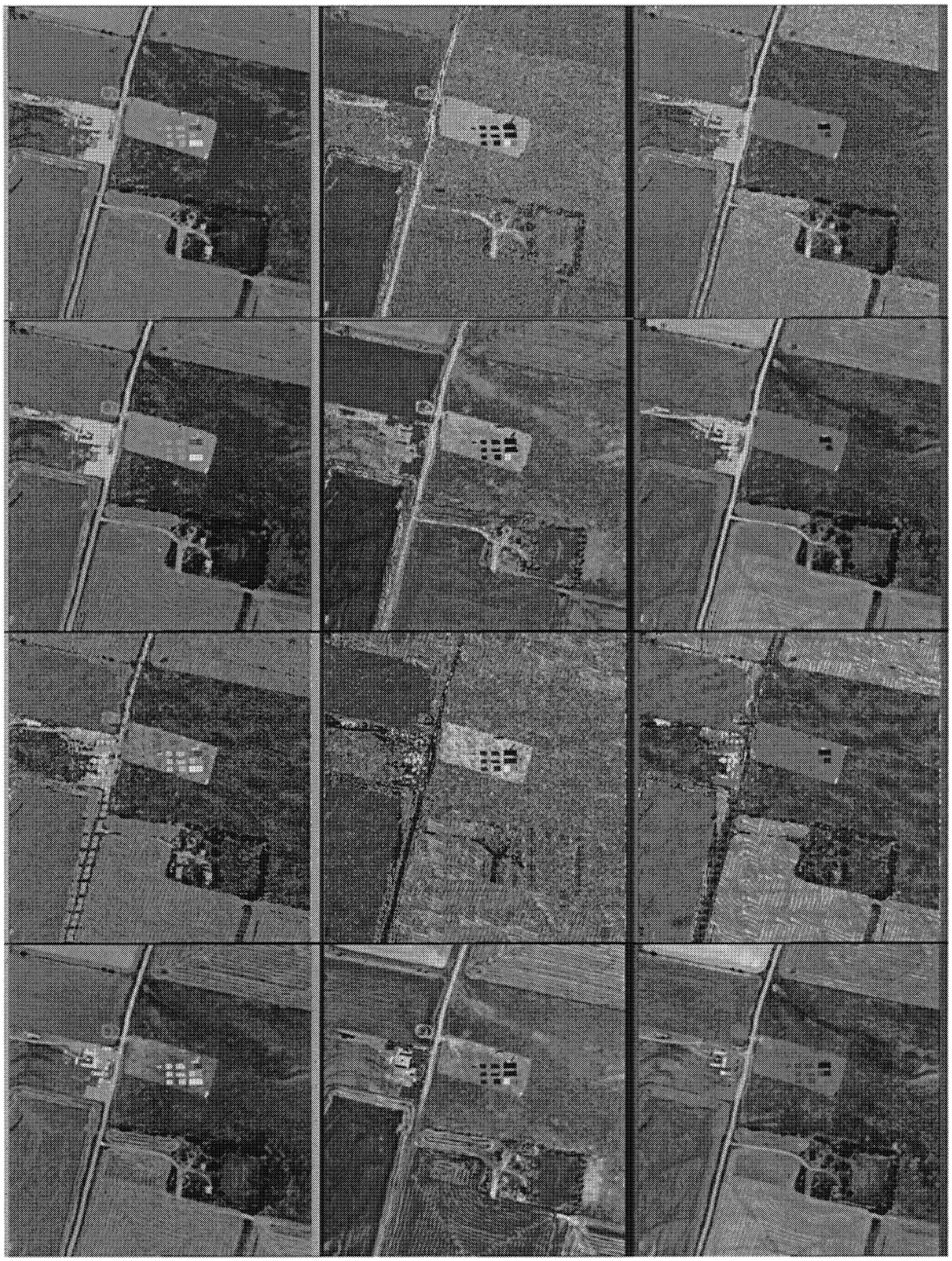


Figure 5(a) – (e): Fig. 5(a) (top row of 3 images) shows 3 sample bands of real HYDICE ARM imagery. Spectral bands 20, 95, and 184 (left to right) with spectral band centers of 0.466, 1.23, and 2.27 microns respectively are shown. Band 20 shows the 9 homogeneous texture test regions, while band 95 shows the 6 transitions regions that have been used for each of the metrics. Fig. 5(b) (second row) shows the DIRSIG image results for the same 3 bands using the SBP texture model. Figs. 5(c), (d), and (e) show the results using the MBP, TS, and FM models respectively.

3.1 SBP Texture Model

The effect of using a single bandpass texture map for rendering an image with broad spectral coverage is evident in Figure 5(b). The later spectral bands all tend toward noise, while the band nearest that of the texture image (band 20, left) has the overall best appearance. This is because the curve selections that were appropriate in the band 20 region were not the correct choices for the non-correlated IR bands. A more detailed discussion is best left until after the MBP model results are presented in the next section.

3.2 MBP Texture Model

The results of using the MBP Z-Score Selection texture characterization model are shown in Figure 5(c). In this case, three bandpasses were used: bands 20 (FWHM = 0.4661 microns), 65 (see Table 3), and 185 (FWHM = 2.2802 microns). There is an obvious improvement made using the MBP model over the SBP model. None of the bands contain noisy structures in the MBP DIRSIG imagery. This is because the composite weighted z-score that is used in the MBP algorithm considers the spectral behavior in multiple bandpasses, and therefore tends to more correctly choose spectral reflectance curves for all pixels in the output image. A precursory visual inspection of especially the later spectral bands of the SBP and MBP results reveals the powerful capabilities of the MBP model over the SBP model, and the quantitative analysis in Section 4 will further demonstrate how much better the MBP model performs for imagery with larger spectral dimension.

3.3 Texture Synthesis (TS) Model

The reader will undoubtedly notice the poorer spatial fidelity of this imagery as compared with the MBP model results. However, it is not as dismal as it may initially appear since the local mean and standard deviation statistics for each material class region in the real and synthetic imagery matched to within 3 %. One must also keep in mind that the very nature of the Texture Synthesis model guarantees that the spectral covariance statistics of the synthetic textures will agree with that of the real image textures. This aspect will be tested in more detail in Section 4 with the SCR and SCM metrics. The uncut pasture and wheat regions are better represented than the plowed field regions, in which the large-scale oriented structural patterns have not been captured. The TS model also had difficulty characterizing within-material transition regions such as the sharp uncut pasture-dirt transition on the right side of the real imagery. One of the fundamental limitations of this model is evident especially in the lower road region. Since the minimum input texture sample size required by the model is 64x64 pixels and the road region is quite narrow, it was difficult to obtain a sample of sufficient size. Therefore, a mirroring utility had to be used in order to grow this region to generate a texture image of sufficient size. This caused repetitive artifacts to appear in the form of dark banding, as shown in Figure 5(d).

3.4 Fraction Map (FM) Texture Model

The FM model produces the most visually pleasing imagery of all four models. In fact, the real and synthetic imagery are almost indistinguishable if one compares Figure 5(e) with the real HYDICE imagery. This re-mixing of fractional abundance planes has produced a very impressive result in the spatial domain and although it appears as though the spectral correlation has been preserved based on a band-by-band visual analysis, the true test will be to determine how well the performance metrics deem this model to be in both spatial and spectral domains in Section 4.

4. TEXTURE MODEL PERFORMANCE ANALYSIS

All synthetic imagery has been rendered such that it is perfectly registered with the real HYDICE ARM data, with identical GSD. This is a strict requirement for the application of the texture model performance metrics since the corresponding pixel data of the real and synthetic imagery have been directly compared in the fifteen texture test regions. Regions 1-9 are homogeneous (within-material class) textures, while regions 10-15 represent transition region textures.

4.1 Spatial GLCM Metric

Due to space limitations, the very detailed performance analysis of the GLCM metric cannot be presented here in its entirety. Further, if the band-by-band analysis had been replaced by synopsis tables consisting of the averaged performance values over the tested spectral band subset then the differential performance between bands would not be captured. Since this presentation of raw data would not be useful, the performance of each model is not presented in table format as in the parent paper to this work[14] but is summarized here in text. The FM model contains the lowest range and average values within the absolute difference images. It also typically possesses the smallest number of outliers from the threshold image. The corresponding pixel values of the absolute difference images are all lower than those of the MBP, SBP, and TS models. This result makes sense intuitively since the spatial appearance of the FM model result is the most visually identical to the real HYDICE ARM image, for all spectral bands. It is therefore clear that the FM model out-performs all of the other models in the spatial domain, for all fifteen of the texture test regions, across all spectral bands. The MBP model results of the GLCM Contrast metric indicate the second best performance spatially. Although the SBP model shows comparable performance for spectral bands 20 and 32, the spatial structure begins to deteriorate for the later spectral bands of the SBP image cube. Therefore, the MBP model performs much better overall, which is also not a surprising result since the MBP model attained the second best ranking through a visual analysis. The SBP and TS models exhibit oscillating behavior in the GLCM Contrast metric result. This is evident even through a visual analysis of the resultant imagery; the earlier spectral bands of the SBP model appear comparable to the result of the MBP model, while the TS model appears to lack the spatial structure present in the real HYDICE imagery. However, the spatial structure is similarly lacking in both the TS and SBP models from band 65 onward. The road region in the TS image tends to suffer spatially due to the artifacts discussed earlier, while that of the SBP model performs somewhat better. The plowed region of the TS model does not capture the oriented structure of the plowed patterns, but nonetheless contains more patterns than the SBP model for the later spectral bands. These results are reflected in regions 1 and 4 (plowed fields) and for regions 12, 14 and 15, which are transition regions including the road. At this point it is suitable to rank the models based on spatial performance in the following manner: 1. FM model; 2. MBP model; 3. SBP model; and 4. TS model. Since the results are very similar in terms of spatial content, the spectral texture analysis of the SCR and SCM metrics will prove to be crucial in distinguishing between the overall performance of the TS and SBP models. The number of outliers from the threshold image for each of the models using the GLCM Correlation metric also indicated the same rank order, although this feature had less discriminatory power than the Contrast metric.

4.2 Spectral SCR Metric

If the spectral structure and complexity is captured in the synthetic image, then its SCR value (for the given region being tested) should theoretically be within an acceptable variance threshold of the corresponding value for the real HYDICE ARM image. The same fifteen texture test regions have been used to obtain SCR values from the SBP, MBP, TS, and FM model DIRSIG images. These values were then compared with the corresponding real image SCR values. The threshold was defined by taking repeated measurements of the SCR from the same regions of the real image that were used to construct acceptable variance threshold images for the GLCM and SCM metrics. This value is the rightmost column of the below table, which indicates the (1 sigma) standard deviation of the repeated SCR measurements from the real image.

Region	SBP	MBP	TS	FM	Real	S.D.
1	56.95	85.36*	91.58*	97.78*	**90.50**	25.94
2	58.28	82.12	95.97*	139.66	**104.74**	15.19
3	76.90	188.47	276.90*	225.29	**411.95**	176.03
4	26.25	68.05*	81.79*	117.19	**77.21**	16.15
5	198.09	106.49*	119.04*	158.92*	**120.25**	39.16
6	464.34	254.69*	200.91*	235.27*	**196.06**	72.86
7	33.19	163.71	297.16*	176.39	**504.28**	237.99
8	773.24	897.42	4,469.35*	943.44	**6,222.54**	2,258.74
9	192.91	274.19	663.82*	297.29	**1.004.99**	486.53
10	103.53	145.25*	147.31*	126.60	**172.51**	43.30
11	110.23	133.75*	142.37*	118.99*	**153.82**	35.32
12	34.48	90.85*	142.68*	162.67	**123.61**	35.86

13	28.94	86.95*	123.45*	195.37*	**119.29**	74.10
14	32.14	73.85*	60.43*	75.01*	**58.90**	18.96
15	145.97	103.46	62.94*	110.35	**72.74**	12.47

Table 1: SCR metric values for SBP, MBP, TS, and FM models for HYDICE ARM imagery. (* = value is within +/- 1 sigma S.D).

The TS model has the closest SCR values to the SCR values of the real HYDICE ARM image. In each case, the TS model image regions have an SCR value within 1 sigma standard deviation of the real corresponding value. This is an intuitive result, since the very nature of the TS model guarantees that the spectral statistics will be correct due to its spectral covariance enforcement methodology in creating synthetic texture. Another consistent observation is that the SBP model has SCR values farthest from the corresponding real image SCR values. In fact, none of its SCR values lie within the standard deviation threshold, although some values are quite close to the threshold. This result is not surprising since the Single-Bandpass z-score curve selection algorithm has only utilized a single narrow band in the visible region of the spectrum, and thus it has tended to select curves for scene materials from the ground truth measurement database that were not optimal for non-correlated spectral bands. The MBP and FM models both perform quite well for this metric; however their relative performance is not consistent. Despite this oscillatory behavior, it is possible to rank one over the other if the results are carefully analyzed. For instance, the MBP model has 9 of its 15 vales within the threshold value, while the FM model has 6. It is also worthy to note that for the three cases in which the FM model SCR values lie outside of the threshold, they were extremely close to the threshold value, as one can infer from the above table. Further, the MBP model attained the second-best performance metric value (next to the TS model) for 10 of the 15 regions, while the FM model did so for the other 5 regions, which incidentally are all homogeneous texture regions. Therefore, since the rankings of each region for the SCR metric are consistent 66.6% of the time, and because the SCR values are so close between the FM and MBP models, the following ranking based solely on the SCR metric result can be concluded: 1. TS model; 2. MBP model; 3. FM model; and 4. SBP model. It is important to note the potential for bias in the SCR results for the TS model since the input texture samples were extracted from the real image in the vicinity of the test region used for SCR measurement. Therefore, the SCR values are certain to be correct for this model. Meanwhile, the performance of the MBP model has great potential for improvement since the image derived "ground truth" spectra used to characterize spectral texture was obtained from much smaller sample regions. Although the use of more complete spectral samples for the MBP model would increase computational requirements, it would likely provide SCR values very close to those of the TS model and of the original image.

4.3 Spatial-Spectral SCM Metric

Both the Contrast and Correlation features of the SCM metric are able to distinguish the performance of each of the texture models quite well in terms of both spatial and spectral structure. As with the other spatial metrics, the best performance values belong to the FM model. This is evident by the average and range of values of each region's absolute difference image, since the entries are much lower than those of the other three models. In order to verify this result (since the average and range are not themselves sufficient to confirm that overall performance is better for this metric), the absolute difference images were compared directly for each model result, and for each texture test region for the Contrast feature of the SCM metric. In all cases, when the FM model metric images were subtracted from the corresponding images of the other models, the result was greater than zero. This indicates that all pixel values were smaller for the FM model, and thus performed the best of all models. Further, the magnitude of the deviation from the threshold was also investigated in order to supplement the information provided by the number of outliers for each texture region. The deviation from the threshold was smallest for the FM model, and second to smallest for the MBP model, despite the fact that the actual percentage of outliers for the FM and MBP models were very close for all 4 spectral band pairs. This indicates that the best overall spatial-spectral performance was achieved by the FM model for the rendering of the HYDICE ARM imagery. It has been shown already that the FM model performs quite well spatially through the application of the GLCM metric. However, since the SCR metric indicated that the spectral performance of the FM model was not as good as for the TS model, and extremely close to the performance of the MBP model, its spatial-spectral overall performance had the potential to suffer. The results of the SCM metric clearly show that by weighting both spatial and spectral dimensions, the FM model performs better than the other models. This means that what the FM models lacks spectrally (compared to the TS model), it makes up for in the spatial domain. Further, the very good results of the SCM metric also shows that the weighting of the end member spectra according to their fractional abundance maps generally creates an adequate level of spectral structure and clutter that is comparable to the real counterpart image. That is, despite the fact that the spectral covariance statistics are not as close to the real image as

the TS model is, the spectral correlation is preserved in the FM model image for this sampling of spectral band pairs, and it is reasonable to infer that this correlation is maintained throughout the spectral extent of the image. This means that the various linear combinations of end member spectra assigned to each mixed pixel is sufficient to represent the spectral clutter present in the real HYDICE ARM image. The number of outliers in the SCM Correlation metric further substantiates the best overall performance of the FM model. The SCM performance metric values that come closest to the FM model belong to the MBP model. There is very clear separation between the MBP model and the FM model, as well as between the MBP model and the SBP and TS models. The range and average values for the MBP imagery are all greater than that of the FM model, but less than those of the TS and SBP models. This is also true for the corresponding pixel values of the absolute difference images; that is, there is no oscillatory behavior between the MBP model and any other model. Although the percentage of outliers of the SCM Contrast metric are relatively close between all models, the magnitude of the deviation from the acceptable variance threshold image is much smaller for the regions of the MBP image than for the SBP and TS models. Therefore, the MBP model has secured the second-best ranking for the SCM metric alone, which is once again reinforced by the results of the number of outliers for the SCM Correlation metric.

The relative performance of the TS and SBP models was the most difficult to discern. For the 95%-correlated band pair (bands 22 and 32), the SBP model performs better than the TS model in general, although the difference tended to be subtle, since the values were lower for the SBP model for 60% of the regions, of which all but one were homogeneous texture regions. However, for the remaining three band pairs, the TS model achieved better performance values. This is an intuitive result since the SBP and TS models shared similar spatial performance metric values for the later spectral bands, while the SBP model performed better spatially for band 20 and often for band 32. Since the SBP model uses only one bandpass for spectral reflectance curve selection, it was able to choose spectra for each pixel that were optimized for that region of the spectrum. It therefore was able to maintain the spatial-spectral correlation between bands 22 and 32 better than the TS model because these bands are so well correlated. This broke down for the band pairs that involved later spectral bands due to the lack of adequate spatial structure in those bands (as observed with the GLCM metric), and due to the poor spectral performance in those bands (as evidenced by the performance of the SBP model with the SCR metric). Since the spectral performance is much better for the TS model than the SBP model, and because the spatial performance for the later spectral bands of the SBP and TS models are comparable, the SCM metric is able to account for both of these aspects and provide metric values that weight spatial and spectral performance simultaneously. In general then, it is reasonable to rank the overall performance of the TS model higher than the SBP model for the HYDICE ARM imagery despite the fact that the SBP model performs slightly better for bands well-correlated with the visible region of the spectrum. Further, for the SCM metric alone (both Contrast and Correlation features), the TS model out-performs the SBP model for 75% of the tested band pairs. Therefore, the ranking based on the SCM metric alone is: 1.FM model; 2. MBP model; 3. TS model; and 4. SBP model. The SCM metric is therefore successful as a simultaneous measure of spatial and spectral fidelity of synthetic image texture, although it tends to weight spatial more than spectral structure since each measure only involves two spectral bands, and not the entire spectral extent of the image. This is an important result since this metric provides the initial mechanics for summarizing the relative performance of all four models for the HYDICE ARM imagery. This is true because the separate analyses of the GLCM and SCR metrics suggested that the FM model always performed the best spatially, and the MBP model performed second-best. Also, the oscillatory behavior observed with the GLCM metric made it very difficult to conclusively rank the relative performance of the SBP and TS models. This was designed to be alleviated by considering the spectral domain as well in order to diagnose the overall performance of each model. The SCR metric verified that the TS model performed extremely well spectrally, while the SBP model performed the worst of all four models. The result was two quite different rank orders from separate spatial and spectral analyses. The question then became, "how much does one weight the spatial and spectral performance metrics?" There are many approaches to an objective weighting of each metric result. One possibility is to apply a weight of 1 to each of the3 metrics and sum the rank orders prescribed by each metric for each texture model so that the lowest number represents the best performance. In this case, the result is: 1. FM model (5); 2. MBP model (6); 3. TS model (8); and 4. SBP model (11). Following the same process, but applying weights of 1, 1, and 0.5 respectively to the GLCM, SCR, and SCM metrics produces an overall ranking of: 1. FM model (4.5); 2. MBP model (5); 3. TS model (6.5); and 4. SBP model (9). In both cases the order is identical, and shows that the performance of the FM and MBP models were very close, while there is more separation between these models and the TS and SBP models. Other ranking techniques have be employed to reflect varying combinations of metric weights and the differential spatial-spectral weighting of the SCM metric, all of which suggest the same overall rank order.

5. CONCLUSIONS

The DIRSIG imagery and associated metric results clearly demonstrate the conditions that must be met and the limitations that must be overcome in order to achieve optimal model performance. Each of the texture models requires accurate and thorough ground truth data in order to realistically capture the spatial and spectral variability of scene material classes, especially for the SBP and MBP models. The advantage of the use of the MBP model over the SBP model has been clearly established. The FM model requires adequate ground truth measurements, distinct fraction maps, and robust end member selection processes in order to confidently assign single end member spectra to each fraction plane. These requirements were met in this research, thereby demonstrating that the simple mixing of end member spectra in accordance with their pixel-by-pixel fractional abundances is able to represent realistic levels of spatial and spectral clutter. Although the TS model did not perform as well as the FM, MBP, and SBP models in the spatial domain, its spectral texture was extremely well characterized due simply to the nature of constraint enforcement of the model. The margin of improvement between the SBP and TS models is larger for the SCM than the GLCM metric, which indicates a better overall performance by the TS model. The overall ranking of texture model performance for both homogeneous and transition region textures is: 1. FM model; 2. MBP model; 3. TS model; and 4. SBP model. In terms of a choice of texture models to use, it depends on the specific application, since each model has varying requirements on availability of input imagery. For instance, while the TS model requires minimum input texture image sizes and is computationally expensive, the SBP and MBP models require ground truth reflectance data that adequately represents the spatial and spectral variability of scene materials. Further, the MBP model requires the availability of a few spectral bands whereas the unmixing process involved with the FM model necessitates accessibility to real image data of higher spectral dimension in order to produce enough fractional abundance maps to represent the spatial and spectral clutter present in the scene to be rendered. The choice of model may also depend on the nature of hyperspectral algorithms to be run on the synthetic imagery. For instance, if SCR is an important measure then one may choose the texture model based on the SCR metric ranking alone, subject to availability of real image data and run time considerations.

6. REFERENCES

1. Schott, J.R., Brown, S.D., Raqueno, R.V., Gross, H.N., Robinson, G., "*An Advanced Synthetic Image Generation Model and its Application to Multi/Hyperspectral Algorithm Development*", Canadian Journal of Remote Sensing, Vol. 15, pp. 99-111, 1999.
2. Ientilucci, E.J., Brown, S.D., "*Advances in Wide-Area Hyperspectral Image Simulation*", Proceedings of the SPIE Conference on Targets and Backgrounds IX: Characterization and Representation, Vol. 5075, ISBN: 0-8194-4934-2, April (2003).
3. Sheltler, B.V., Mergens, D., Chang, C., Mertz, F., Schott, J.R., Brown, S.D., *et al.*, "*Comprehensive Hyperspectral Image Simulation I:Integrated Sensor Scene Modeling and Sensor Architecture*", Proceedings of SPIE AeroSense, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI, Orlando FL, Vol. 4049, pp. 94-104, April (2000).
4. Simoncelli, E., Portilla, J., "*Texture Characterization via Joint Statistics of Wavelet Coefficient Magnitudes*", Fifth IEEE International Conference on Image Processing, Chicago, October 4-7, Vol. 1, IEEE Computer Society (1998).
5. Schott, J.R., Salvaggio, C.N., Brown, S.D., Rose, R.A., "*Incorporation of Texture in Multispectral Synthetic Image Generation Tools*", SPIE Aerosense, Targets and Backgrounds: Characterization and Representation, Vol. 2469, No. 23, Orlando, FL, 1995.
6. Kennedy, C.S., "*The Testing and Assessment of Texturing Tools Used to Build Scenes in DIRSIG*", Bachelors' Degree Thesis, Rochester Institute of Technology, Rochester, NY, June (2002).
7. Tyrrell, A., *Multi/Hyperspectral Texture Synthesis*, Masters' Degree Thesis, Rochester Institute of Technology, Rochester, NY, January (2003).
8. Brown, S.D., Ientilucci, E.J., Raqueno, R.V., Schott, J.R., "*DIRSIG/GENESSIS Hybrid Hyperspectral System Simulation of the Fort AP Hill Site*", prepared for Photon Research Associates RIT/DIRS Report 00/01-71-162, Rochester, NY, January (2001).
9. Haralick, R.M., "*Statistical and Structural Approaches to Texture*", Proceedings of IEEE, Vol. 67, No. 5, May 1979, pp.786-804.
10. Gotlieb, C., Kreyszig, H., "*Texture Dexcriptors Based on Co-Occurrence Matrices*", Computer Vision, Graphics and Image Processing, Vol. 51, pp. 70-86.
11. Clausi, D.A., Zhao, Y., "Rapid Extraction of Image Texture by Co-Occurrence Using a Hybrid Data Structure", Computers & Geosciences, Vol. 28, pp. 763-774.
12. Manolakis, D., Shaw, G., Keshava, N., "*Comparative Analysis of Hyperspectral Adaptive Matched Filter Detectors*", Algorithms for Multispectral and Hyperspectral Imagery VI, S.S. Chen and M.R. Descour eds., Vol. 4049, pp. 2-17, SPIE Orlando FL, April (2000).
13. Hauta-Kasari, M., Parkkinen, J., Jaaskelainen, T., Lenz, R., "*Multispectral Texture Segmentation Based on the Spectral Co-Occurrence Matrix*", Pattern Analysis & Applications, Vol. 2, pp. 275-284.
14. Scanlan, N.W., *Comparative Performance Analysis of Texture Characterization Models in DIRSIG*, Masters' Degree Thesis, Rochester Institute of Technology, Rochester, NY, 2003.