AN ALGORITHM FOR GENERATING NATURAL COLOR IMAGES FROM FALSE COLOR USING SPECTRAL TRANSFORMATION TECHNIQUE WITH HIGHER POLYNOMIAL ORDER

Sumi Kala*, Aravinda Kumar, A. K. Joshi, Vinod M Bothale, B Gopala Krishna

DPPA&WAA, National Remote Sensing Centre, Hyderabad – (sumi_kala, aravindakumar_p, joshi_ak, vinod_bothale, bgk)@nrsc.gov.in

Commission V, SS: Emerging Trends in Remote Sensing

Satellite imageries in True color composite or Natural Color composite (NCC) serves the best combination for visual interpretation. Red, Green and Infrared channels form false color composite which might not be as useful as NCC to a non-remote sensing professional. As blue band is affected by large atmospheric scattering, satellites like IRS-LISS IV, SPOT do not have blue band. To generate NCC from such satellite data blue band must be simulated. Existing algorithms of spectral transformation do not provide robust coefficients leading to wrong NCC colors especially in water bodies. To achieve more robust coefficients, we have proposed new algorithm to generate NCC for IRS-LISS IV data using second order polynomial regression technique. Second order polynomial transformation functions consider even minor variability present in the image as compared to 1st order so that the derived coefficients are adjustable to accommodate spatial and temporal variability while generating NCC. In this study, Sentinel-2 image was used for deriving coefficients with blue band as dependent and green, red and infrared as independent variables. Simulated Sentinel band showed high accuracy with correlation of 0.93 and 0.97 for two test sites. Using the same coefficients, blue band was simulated for LISS-IV which also showed good correlation of 0.90 with sentinel original blue band. On comparing LISS-IV simulated NCC with simulated NCC from other algorithms, it was observed that higher order polynomial transformation was able to achieve higher accuracy especially for water bodies where expected color is green. Thus, proposed algorithms can be used for   transforming false color image to natural color images.

1. Introduction

Natural Color Composite images represent features in their true color and have been widely used in visual interpretation, display or fly-through of draped terrain. False Color Composite is generated using near-infrared, red and green bands which is not natural to human eye. This does not have any problem in extracting information for person trained with remote sensing techniques however, for non-remote sensing professional NCC is superior to FCC as they are not comfortable with false color look (Patra et al., 2006). High resolution true color image could be on the most comfortable one for conventional interpretation and visualization.

The sun emits full spectrum but only a small part of energy reaches earth's surface because of scattering. The shorter wavelengths are scattered more because of the presence of small atmospheric particles. Blue wave suffers maximum scattering in the visible region, approximately 5.5 times greater than red and green wave (Zhu et al.2012). Therefore, image in blue wavelength tends to appear hazier in comparison to other bands. Consequently, some of the multi-spectral sensors do not include blue band for example, Indian Remote Sensing (IRS)-P6 sensor covers only two bands in visible region i.e. green and red one in NIR region.

Recent advances in online geospatial portal like Bhuvan, Google Maps, especially for nonprofessionals, have led to the urgent demand for NCC images with high spatial resolutions. LISS -IV provides high spatial resolution data of 5.8 meters however it cannot form natural color composite image from available bands. Hence, to achieve Natural color, blue band requires to be being simulated using spectral transformation techniques.

Several authors have carried out various studies for simulating blue band. Chen and Tsai (1998) proposed spectral-transformation method in order to transform SPOT FCC into NCC. It was based on the selection of spectral control points from Landsat and SPOT images and the first-order polynomial transformation function. Cardei (1999), studied four methods i.e. averaging, linear regression, polynomial regression and neural network for recovering blue band from two bands i.e. green and red. He concluded that complex method like neural network does not drastically improve the accuracy in comparison to less complex method linear regression.

Knudsen, 2005 developed an algorithm using color infrared aerial photo based on urban/suburban surrounding. It consisted of least-squares adjusted linear model establishing relationship between blue band and red, green, near-infrared channels. However, the Least Square Regression model was very simplified and recommended to develop new regression

* Corresponding Author

for different data sets. Patra et al., 2006 used a similar approach of spectral transformation to establish a relationship between FCC and NCC using IKONOS data having a broader spectral coverage. A set of adjustable radiometric transformation coefficients were derived to accommodate variation in spatial and dynamic range offered by sensors to generate natural color. The disadvantage of 1st order polynomial transformation function is that it does not provide robust coefficients. These coefficients are not able to accommodate the variability in the scene. Hence, giving wrong color especially for water bodies.

Xu Huixi and Chen Yunhao, 2012 developed a Technique for Simulating Pseudo Natural Color Images Based on Spectral Similarity Scales. This technique is similar to look up table which is used to find the most similar spectrum in spectral library Smaller value corresponds to spectra that is most matched and then minimum SSV best similarity spectrum can be searched in spectral library. Zhu et al. 2012, developed a non-linear model based on spectral library and machine learning to simulate SPOT NCC images from FCC. It required first to train spectral samples using machine learning approach such as SVM or Back propagation neural network. The resulting relationship is used to construct a sophisticated nonlinear model. This limitation of these techniques was that it does not take into account the mixed-pixel effect.

Some of existing methods discussed in above paragraphs can be describe in brief as follows:

(i) Averaging Methods: This is a simple approach to derive NCC in which (i) original green band is assigned to blue band and red remains as it is (ii) green band is then simulated by average band computation (Cardie, 1999 , Bulatovic et al, 2013). ERDAS Imagine Software uses a similar weighted averaging approach for simulating NCC which can be expressed as:

$$B.c = G$$
$$G.c = \frac{3\,G + NIR}{4} \qquad (1)$$
$$R.c = R$$

Where Bc, Gc, Rc, NIRc= simulated Blue, Green, Red and Near Infrared band, B, G, R, NIR= original Blue, Green, Red and Near Infrared band,

(i) Regression: This approach involves spectral transformation based on linear and nonlinear regression function.

    a. Linear Regression: It models a linear relationship between the dependent and independent variables mostly using least squared model such that residual squared error for the calibration data is minimized (Knudsen, 2005, Zhu et al, 2012). The objective is to get the 'best fit' values of the coefficients m1, m2 and m3 by least :

$$B = {}_{m1}NIR + {}_{m2}R + {}_{m3}G \qquad (2)$$

where, NIR, R, G and B are column vectors containing the reflectance of the respective channels and m1,m2,m3 are coefficients.

    b. Polynomial Regression: Polynomial models are used to approximate the nonlinear relationship between dependent and independent variable. Arbitrary fitting of higher order polynomial can cause severe over fitting. Chen and Tsai, 1998 developed spectral transformation algorithm by selecting few control points using a first order polynomial function between Landsat TM and SPOT images followed by derivation of transformation coefficient using least squared regression.

(ii) Neural Network: Neural Network algorithms, initially proposed by Cardei, 1999 describes the complex non-linear relationships between Blue and other channels. Later Zhu et al, 2012 used back propagation artificial neural network for simulation blue band by training original blue band.

The objective of the study was to develop a relation between blue and green, red and near infrared bands using second order polynomial regression transformation. Sentinel-2 multispectral data is used as reference image for deriving robust coefficients which were further used to simulate blue band for IRS LISS-IV imagery and achieve pseudo natural color composite.

2. Data Used

2.1 Sentinel-2: It covers visible to SWIR region of electromagnetic spectrum consisting of 13 spectral bands with different spatial resolutions ranging from 10m to 60m. In this study we selected Blue (Band 2), Green (Band 3), Red (Band 4) and NIR (Band-8) for simulation. For the purpose of deriving coefficients, data was clipped to small region consisting of all land classes including vegetation, barren land and water body.

Wavelength (μm)	
Sentinel-2 (10 m)	LISS-IV (5.8 m)
Band 2 : 0.490	Band 2 (0.52-0.59)
Band 3: 0.560	Band 3(0.62-0.68)
Band 4: 0.665	Band 4 (0.77-0.86)
Band 8: 0.842	

Table 1: Wavelength of Sentinel-2 and IRS LISS-IV

2.2. IRS LISS IV: LISS IV has three spectral bands with 5.8 meter spatial resolution. Since, it does not have blue band coefficients derived from Sentinel were applied to this data to simulate pseudo NCC. A brief description of the band specifications is given in Table 1.

3. Methodology

3.1 Top of Atmospheric Correction (TOA)

LISS-IV image is provided with DN whereas; Sentinel is download L1A which is TOA corrected. Therefore, LISS-IV DN was converted to radiance to TOA reflectance using the following equations:

$$\rho = \frac{\pi . L_\lambda . d^2}{EUSN_\lambda . \cos(\theta)} \qquad (3)$$

$$L_\lambda = Gain * DN + Bias \qquad (4)$$

Where ρ is the planetary TOA reflectance, L_λ is the spectral radiance at the sensor's aperture, d is the earth-to-sun distance, $ESUN_\lambda$ is the mean exo-atmospheric solar irradiance, θ is the solar zenith angle (in degrees), Gain is the band-specific rescaling gain factor and Bias is the band-specific rescaling bias factor. The proposed methodology is summarized in Figure 1.
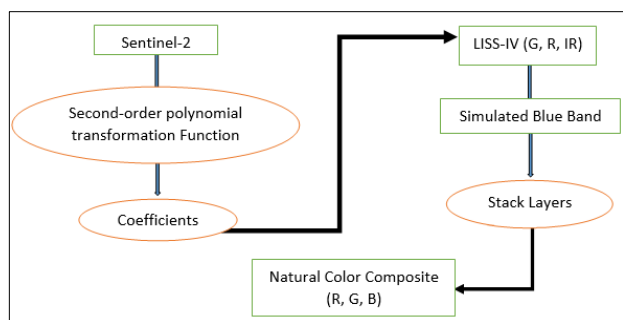

Figure 1: Flow chart of methodology adopted in this research.

### 3.2 Second Order Spectral Polynomial Transformation

Sentinel -2 image was selected to derive coefficients using polynomial transformation. The purpose is to capture the trend in the data by assigning a single function across the entire range. Each of the four bands was converted separately to column vectors for further processing. 1st order of polynomial does not accommodate to the variability in the scene therefore, 2nd order polynomial is selected. Figure 2 shows scatterplot of blue band against near infrared band and the trend line using linear and polynomial fitting for 40 pure pixels taken randomly over the study area.

The 2nd order polynomial transformation with three independent variables (Green, Red and Infrared) and one dependent variable (Blue) is given by
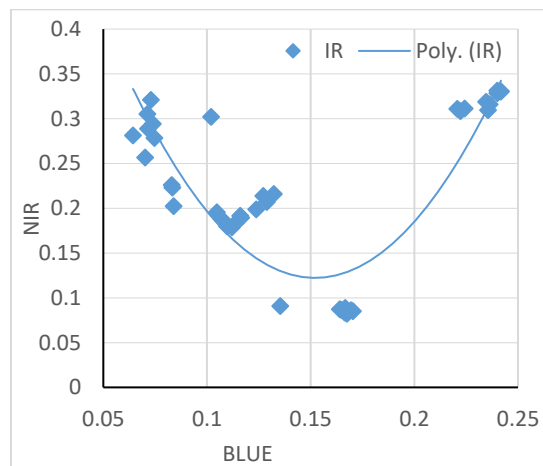
$$Bs = [G, R, IR], B \qquad (5)$$

$$B' = \beta_o + \beta_1 * G + \beta_2 * R + \beta_3 * IR + \beta_4 * G^2 + \beta_5 * R^2 + \beta_6 * IR^2 + \beta_7 * G * R + \beta_8 * G * IR + \beta_9 * R * IR + \beta_{10} * G * R * IR \qquad (6)$$

where, βn are the polynomial coefficients, B'= simulated blue band, B= Blue, G= Green, R= Red and NIR= Near Infrared band from Sentinel-2 data.
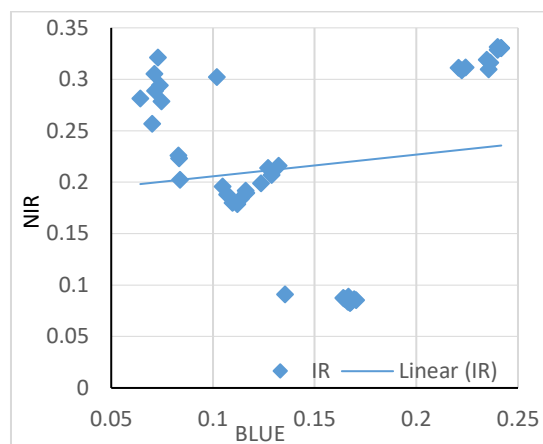
(a)

(b)

Figure 2: Scatterplot (NIR vs. Blue) (a) Linear fitting (b) Polynomial Fitting

The polynomial regression model can be developed to achieve the best fit curve using the least squares methods which aim to minimize the variance between the values estimated from the polynomial and the values from the dataset. The 'best' line is equivalent to minimize the sum of squared error between the line and data points. In order to statistically verify the coefficients root mean squared error, sum of squared error and correlation coefficient were derived.

3.3 Natural Color Simulation

A set of 11 coefficients were derived from the polynomial transformation which were used to derive simulated blue band for Sentinel-2 and LISS-IV based on equation 6. Simulated blue band is reshaped from column vector into

original row * column format and stacked with original green and red bands to form pseudo natural color composite

4. Results and Discussion

In order to confirm that the proposed algorithm is feasible for simulating NCC, two test sites were selected for Sentinel data which were from different time and land cover. For LISS-IV, a comparison was made by simulating NCC from proposed and existing algorithms like ERDAS and Linear Transformation. The results were examined statistically as well as visually.

4.1 Sentinel-2: Natural Color Simulation:

Figure 3 and Figure 4 shows the visual comparison of True Color Composite and Simulated Natural Color Composite of Sentinel-2 for different land cover types. Figure 3 highlights test site 1 consisting of vegetation, water body and fallow land whereas Figure 4 highlights test site 2 consisting of urban area and also it belongs to a different acquisition date.

Figure 3 shows Sentinel Test Site 1 (a) True Color Composite (b) Simulated Natural Color Composite

Figure 4 shows Sentinel Test Site 2 (a) True Color Composite (b) Simulated Natural Color Composite

	Test Site 1		Test Site 2	
	Blue	Blue'	Blue	Blue'
Min	0	0	0.08	0.06
Max	0.25	0.22	0.41	0.42
Mean	0.1066	0.1081	0.1195	0.1197
S.D	0.0140	0.0153	0.020	0.026
R^2	0.93		0.97	

Table 2: Statistical results for the difference between blue channel and estimated blue channel

Water body, vegetation and urban land use land cover classes in simulated NCC have high visual similarity to that of true color composite however, to eliminate the subjectivity in results statistical comparison were made by estimating the minimum, maximum, mean, median and correlation coefficient between original blue band and simulated blue band for both the areas. Table 2 shows the comparison of original and simulated blue band. In both test sites, correlation coefficient was found to be very high i.e. 0.93 and 0.97 for test site 1 and 2 respectively indicating that

the simulated blue bands have good agreement with original blue bands. Comparatively lower values of correlation in Test site 1 indicates that pixel values of water body are quiet distinct in NIR band which do not fall as close to best fit line thus limiting the prediction of pixel values for water bodies to a certain extent. Overall, it can be infer from results that the algorithm can be used to recover the Sentinel-2 blue band with high levels of accuracy.

In addition to statistical significance, it is important to ensure that each class conserves its spectral properties so some training sites were randomly collected from each class i.e. vegetation, soil and water and their means spectral signature was compared with their true signature as shown in Figure 5. Band 2 and band 3 were used from the original band hence no variation in spectral profile is shown. Also, spatial profile X-X' (Figure 3) was drawn to analyze the spatial variability of simulated blue band which shows high similarity with original blue band (Figure 6). It was observed that algorithm was able to maintain spatial as well as spectral characteristics with high accuracy.
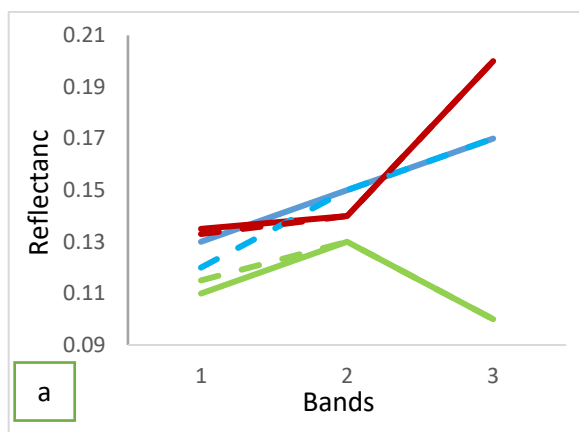
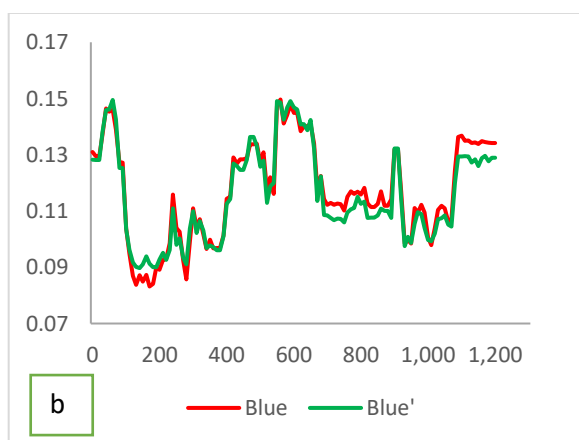| NCC	Vegetation	Water	Soil
Original			
Simulated			

Figure 5: a. Spectral profile (b) Spatial Profile across part of Test Site 1

4.2 LISS-IV Natural Color simulation

The main objective of this study was to derive a set of robust coefficients that can adjust the variability in time and space and can be utilized for simulating blue band for LISS-IV. Figure 6 shows LISS-IV simulated

NCC of Test Site 1 where water body, vegetation and land shows high similarity in visual appearance with that of Figure 3a. Visual analysis on quality of spectral transformation could be highly subjective and also LISS-IV does not have blue band to compare the simulated results. Therefore, statistical estimates are made by comparing simulated blue band with Sentinel blue band which yields a good correlation of 0.90.

Figure 6: LISS-IV (Test Site 1) Simulated Natural Color Composite

This result was also compared NCC simulated using Linear Transformation and ERDAS NCC algorithm.

Figure 7: NCC simulated from Linear Transformation

a. Linear Transformation is able to achieve true color for vegetation and other land cover except water body as

shown in Figure 7. This is because the blue band does not have linear relationship with near infrared band unlike green and red bands.

b. ERDAS replaces blue band with green and simulates green band with averaging methods. However, it is not able to achieve accurate true colors. Vegetation appears to be unnatural green and water body appears in darker shades of blue where green is expected (Figure 8).

Figure 8: NCC simulated from ERDAS Natural Color algorithm

5. Conclusion

In this study, we have developed a second order spectral transformation method to transform false color images into natural color images. The proposed method derives a set of transformation coefficients which can be used to simulate blue band for Sentinel as well as LISS-IV data. The methodology gave high accuracy results for images taken from region and classes. Advancement can be made so as to further maximum the achievable accuracy for different season data with higher variability.

References

Chen, C.F. and Tsai, H.T., 1998. A spectral transformation technique for generating SPOT natural colour image. *In Poster Session 3, GIS Development Proceedings, ACRS*.

Patra, S.K., Shekher, M., Solanki, S.S., Ramachandran, R., and Krishnan.R, 2006. A technique for generating natural colour images from false colour composite images. *International Journal of Remote Sensing*, 27(14), pp. 2977-2989.

Huixi, X. and Yunhao. C., 2011. A technique for simulating pseudo natural color images based on spectral similarity scales. *Geoscience and Remote Sensing Letters, IEEE*, 9(1) pp. 70-74.

Bulatovic, V., Vasic, D., and Ninkov, T., 2013. Method for generating natural color from false color images based on Normalized Difference Vegetation Index clustering

Knudsen, T., 2005. Pseudo natural colour aerial imagery for urban and suburban mapping. *International Journal of Remote Sensing*, 26(12), pp. 2689-2698.

Cardei, C.V. From Dichromatic to Trichromatic Images; *In Proceedings of the Image Processing, Image Quality, Image Capture Systems Conference (PICS '99),* April 25–28, 1999, Savannah, Georgia, 359–363 (Springfield, VA, USA: The Society for Imaging Science and Technology).ions for Image Recovery and Visualization.

Zhu, C., Luo, J., Ming D.,, Shen, Z., and Li, J., 2012. Method for generating SPOT natural-color composite images based on spectrum machine learning. *International Journal of Remote Sensing*, 33(4), pp. 1309-1324, doi:10.1080/01431161.2010.550652. 2012