Spectral Inpainting for the Restoration of Missing Data from Multispectral Satellite Sensors: Case study on Aqua MODIS Band 6

Marouan Bouali, Saïd Ladjal

International Journal of Remote Sensing
Vol. 00, No. 00, DD Month 200x, 1–10

Research Article

Spectral Inpainting for the Restoration of Missing Data from Multispectral Satellite Sensors: Case study on Aqua MODIS Band 6

Marouan Bouali*† and Saïd Ladjal†

† Télécom ParisTech, Département Traitement du Signal et des Images, CNRS LTCI, 75634 Paris Cedex 13

(*v3.6 released October 2010*)

In this letter, we introduce an algorithm for the restoration of missing data from multispectral satellite imagery. The proposed approach combines two simple principles; non local or neighborhood filters used in the context of still image denoising/inpainting and spectral matching techniques based on spectral similarity measures and required for the classification of hyperspectral images. The resulting semi-physical approach, refered to as spectral inpainting, is applied to the particular issue of Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) band 6 (1628-1652 μm) non functional detectors and is shown to provide very good results compared to standard interpolation techniques.

1. Introduction

The post-launch phase of a satellite imaging instrument can reveal dramatic deviations from the initial characterization of its spectro-radiometric response. This is indeed the main motivation behind the use of on-orbit calibration components, which provide complementary tuning parameters for the acquisition system. Unfortunaly, technological alternatives are yet to be conceived when specifics parts of the sensor become completely inoperable, either transiently or permanently. A common issue that occurs in such cases is the loss of information, visible in acquired images as missing pixels. Major factors contributing to data loss include for example downlink transmission errors between the satellite and the receiving groud station, inwhich case partial retrieval of information leads to anomalous reconstruction of images. Another cause for missing pixels reported on several imaging spectrometers such as Landsat, SPOT and MODIS is detector failure due to temporary space particle bombardement or irreversible hardware damage. Inoperable detectors can be highly problematic for both pushbroom and whiskbroom scanners given their role in the image acquisition process. In fact, one dimension of the image is generated from the orbital motion of the satellite while the other dimension is captured by individual detectors. Consequently, a dead-detector will result in a single line drop-out for pushbroom sensors and periodic missing lines from the entire swath for whiskbroom instruments. A good example of the later scenario occurs in Aqua MODIS band 6 (1628-1652 μm). NASA's MODIS instrument is a cross-track scanner composed of a double sided scanning mirror that deflects the

*Corresponding author. Email: marouan.bouali@inria.fr

International Journal of Remote Sensing
ISSN 0143-1161 print / ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/0143116YYxxxxxxxx

Earth energy into 4 Focal Plane Assemblies (FPA). Each FPA is sensitive to a specific portion of the electromagnetic spectrum and contains detector arrays of 10, 20 or 40 pixels depending on the resolution of the considered spectral band. On Aqua MODIS band 6, 8 of the 20 detectors are non functional and do not capture any signal. As a result, data from Aqua MODIS band 6 is affected with a striping pattern that compromises visual interpretation and discourages any further quantitative processing. Furthermore, similarly to many other spectral bands, radiometric miscalibration of detectors also generates stripe noise. The spectral location of band 6 discards its usefulness for the estimation of forest biomass (Baccini *et al.* 2004), aerosols concentration and optical depth (Ignatov *et al.* 2005) and more importantly cloud detection and snow cover (Salomonson and Appel 2004). Given the high necessity of snow cover mapping and monitoring in studies related to climate change, many scientists have relied instead on the use of Aqua MODIS band 7 (2105-2155 μm) to compute the Normalized Difference Snow Index (NDSI). This alternative however decreases the accuracy of retrieved NDSI measurements and only applies to snow applications because of similar snow reflectance values in bands 6 and 7. Following these observations, two methods have been proposed to restore Aqua MODIS band 6. Both exploit the strong correlation between band 6 and band 7. In (Wang *et al.* 2006), missing pixels are estimated via global interpolation. A polynomial regression between Terra MODIS bands 6 and 7 swaths is established and transposed to Aqua MODIS. Over snow covered areas, the authors suggest the following analytical relationships:

$$\begin{aligned} \rho_6 &= 1.6032\rho_7^3 - 1.9458\rho_7^2 + 1.7948\rho_7 + 0.012396 \\ \rho_6 &= -0.70472\rho_7^2 + 1.5369\rho_7 + 0.025409 \end{aligned} \tag{1}$$

where ρ_6 and ρ_7 are Top-Of-Atmosphere (TOA) reflectances for bands 6 and 7. As pointed out in (Wang *et al.* 2006), global interpolation does not account for scene cover types and restored results over ocean, land or desert are often disappointing. This limitation is partly accounted for in the local interpolation approach introduced in (Rakwatin *et al.* 2009). The described algorithm uses a sliding window that locally evaluates the cubic polynomial relationship between bands 6 and 7 valid pixels. This methodology however does not necessarily ensure cover type distinction because its locality is restrained to the spatial dimension. In fact, a missing pixel from band 6 is derived from its value in band 7, when the later does not correspond to a minima or maxima of the considered window. If this implicit assumption of data homogeneity is not satisfied, missing data is estimated from pixels spatially close but spectrally distant.

In this letter, we introduce a general concept to estimate the value of dead pixels from multispectral imagery. The method relies on a simple, yet physically coherent approach that combines non local filters and spectral similarity, respectively used in image restoration and hyperspectral image classification.

2. Methodology

Let us consider the general case of a multispectral intrument composed of K spectral bands with $K \geq 3$ which corresponds to sensors that collect data from at least the red, green and blue regions of the electromagnetic spectrum. To simplify notations in the following, we assume that only one band, denoted as band number M is suffering from missing data. The signal acquired by the imaging scanner is considered as a vector in a K-dimension space, defined at pixel x of a bounded

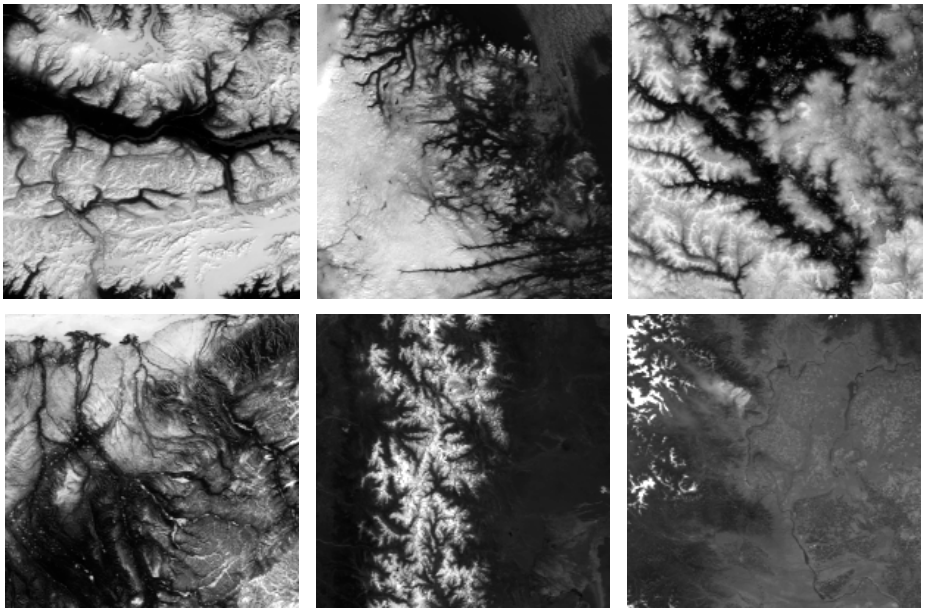


Figure 1. Images from Terra MODIS Band 3 acquired over snow covered areas and used to validate the restoration of Aqua MODIS band 6 (a) Alaska (b) Labrador (c) Siberia (d) Kuparuk (e) South America (f) Washington

domain $\Omega \in \mathbb{R}^2$ as:

$$\rho(x) = (\rho_1(x), \rho_2(x)..., \rho_i(x), ..., \rho_K(x))^T \tag{2}$$

where $\rho_i(x)$ is the reflectance of pixel x measured in the spectral band i.

2.1 *Spectral Similarity*

In order to ensure physical consistency in the restoration of band M, the spectral characteristics of missing pixels are required. This however cannot be achieved using only information from band 6, and indeed explains the poor results obtained with standard interpolation techniques that consist for example in replacing non-functional detectors with neighbors. It should be clear that the analysis of additional spectral bands offers the possibility to discriminate different cover types. To this purpose, we naturally rely on spectral similarity measures commonly used for the segmentation and classification of hyperspectral remotely sensed data. For instance, the determination of surface composition in most geological disciplines, lies upon quantitative comparison between the reflectance values of acquired pixels and pre-determined reflectance spectra. Such spectral matching approach exploits spectral similarity measures (SSM) to distribute pixels withing separate classes. Let us recall the definition of few SSM often used in imaging spectrometry. The Spectral Angle Measure (SAM) is defined in (Kruse *et al.* 1993) as the following angle:

$$\mathrm{SAM}(x, y) = \arccos\left(\frac{\sum_{i=1}^{K} \rho_i(x).\rho_i(y)}{\sqrt{\sum_{i=1}^{K} \rho_i^2(x) \sum_{i=1}^{K} \rho_i^2(y)}}\right) \tag{3}$$

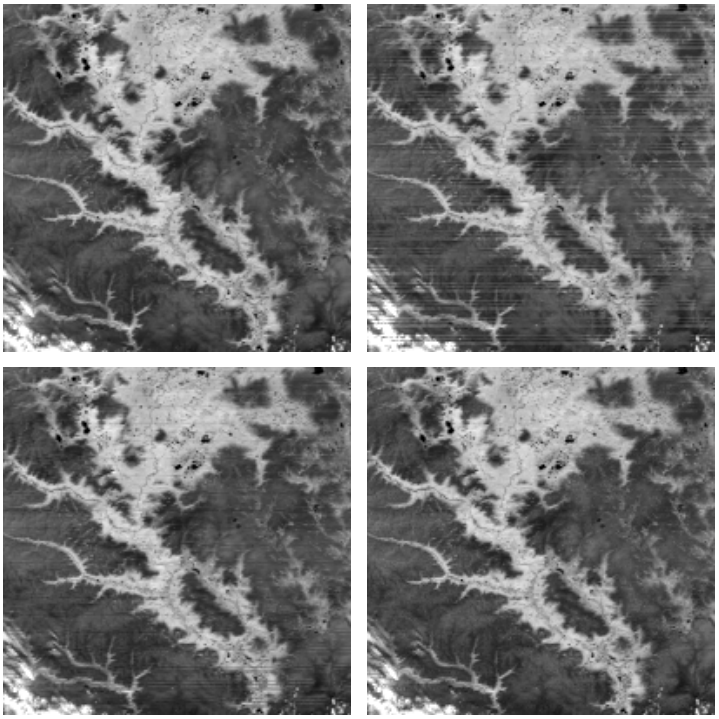


Figure 2. (a) Original image from Terra MODIS band 6 (Siberia) (b) Restored with global interpolation (Wang et al.) (c) Restored with local interpolation (Rakwatin et al.) (d) Restored with proposed spectral inpainting using the EDM

The Euclidean Distance Measure (EDM) between two pixels x and y is given by:

$$\mathrm{EDM}(x,y) = \sqrt{\sum_{i=1}^{K} \left(\rho_i(x) - \rho_i(y)\right)^2} \tag{4}$$

The Spectral Information Divergence Measure (SIDM) is a stochastic index that measures the distance between the probability distribution of spectrum associated with pixels x and y:

$$\mathrm{SIDM}(x,y) = D(x||y) + D(y||x) \tag{5}$$

$D(x||y)$ is the relative entropy of y with respect to x computed as:

$$D(x||y) = \sum_{i=1}^{K} p_i(x) D_i(\rho_i(x)||\rho_i(y)) = \sum_{i=1}^{K} p_i(x)(I(\rho_i(x)) - I(\rho_i(y))) \tag{6}$$

where

$$p_i(x) = \frac{\rho_i(x)}{\sum_{j=1}^{n} \rho_j(x)} \text{ and } I(\rho_i(x)) = -\log\, p_i(x) \tag{7}$$

For a detailled discussion on the effectiveness of these measures in the case of hyperspectral imagery, we refer the reader to (Meer 2006).

Table 1. Information about Terra MODIS swaths used in this study

Region	Day	MODIS Product
Alaska	05/12/01	MOD021KM.A2001132.2115.005.2010063115234
Labrador	11/07/00	MOD021KM.A2000312.1545.005.2010049235331
Siberia	05/24/01	MOD021KM.A2001144.0510.005.2010064034702
Kuparuk	05/23/02	MOD021KM.A2002143.2155.005.2010083121102
South America	12/08/01	MOD021KM.A2001342.1445.005.2010075210602
Washington	07/22/02	MOD021KM.A2002203.1905.005.2010087151736

2.2 *Nonlocal inpainting*

The restoration of missing data is a popular topic in the area of still image processing, often refered to as inpainting. Unlike satellite imagery, the primary goal is purely cosmetic and damaged areas are recovered in a way that enables transparency of the initial degradations. As such, any considered restoration procedure aims to garantee continuation of structures and/or texture in the restored image. Among all techniques developped for the inpainting issue, we draw a particular attention to the seminal work of (Efros and Leung 1999). The algorithm introduced by the authors exploits redundancy and self similarity to synthetise texture from a given sample. Square patches of predetermined size in the sample image are compared to the neighborhood of a pixel in terms of L^2 distance. A random selection among matching patches is used in a copy/paste procedure that gradually fills in the missing area starting with pixels along its border.
The very concept of nonlocal self-similarity has also been exploited for denoising purposes. The Yaroslavsky filter for example (Yaroslavsky and Eden 1996), denoises data using an average of pixels with similar grey-level intensity. Perhaps closer in principle to the Efros and Lung approach, is the recent NL-means denoising algorithm proposed in (Buades *et al.* 2005), where an entire local configuration centered on a noisy pixel is considered. In both cases, an estimate $\hat{u}(x)$ of the true image at pixel x is given by:

$$\hat{u}(x) = \frac{1}{C(x)} \int_{\Omega} w(x,y) f(y) dy \tag{8}$$

where f is the noisy image, C a normalizing parameter and $w(x,y)$ is a weighting function that quantifies the similarity between pixels x and y. For the Yaroslavsky filter, the weighting function w can be expressed as:

$$w(x,y) = exp\left(-\frac{|f(x) - f(y)|^2}{h^2}\right) \tag{9}$$

where h is a parameter that controls the decay of weighting coefficients.
Going back to the issue of multispectral imagery restoration, we argue that the information of the spatial dimension is of little interest compared to the one contained in the spectral axis. Suffice it to notice the analogy between neighborhood similarity of nonlocal filters and spectral similarity used in satellite image classification to invoque the proposed concept of *spectral inpainting*. Let us consider a Spectral Similarity Measure (SSM) from those defined in the previous section (SAM, SIDM and EDM). Since we assumed that spectral band M is corrupted, it should not be included in the computation of a SSM. For example, the EDM

between pixels x and y is computed as:

$$EDM(x,y) = \sqrt{\sum_{i=1,\ i\neq M}^{K} \left(\rho_i(x) - \rho_i(y)\right)^2} \tag{10}$$

and the spectral neighborhood of a pixel x is defined as:

$$S(x) = \{x \in \Omega | SSM(x,y) < \epsilon \text{ and } \rho_M(x) \neq 0\} \tag{11}$$

where the condition $\rho_M(x) \neq 0$ discards missing pixels of band M and ϵ is a spectral deviation tolerance parameter. Following the same approach used by nonlocal denoising algorithms, the value of a missing pixel x from band M can be estimated from:

$$\hat{\rho_M}(x) = \frac{1}{C(x)} \int_{S(x)} \rho_M(y) e^{\frac{-SSM(x,y)}{h^2}} dy \tag{12}$$

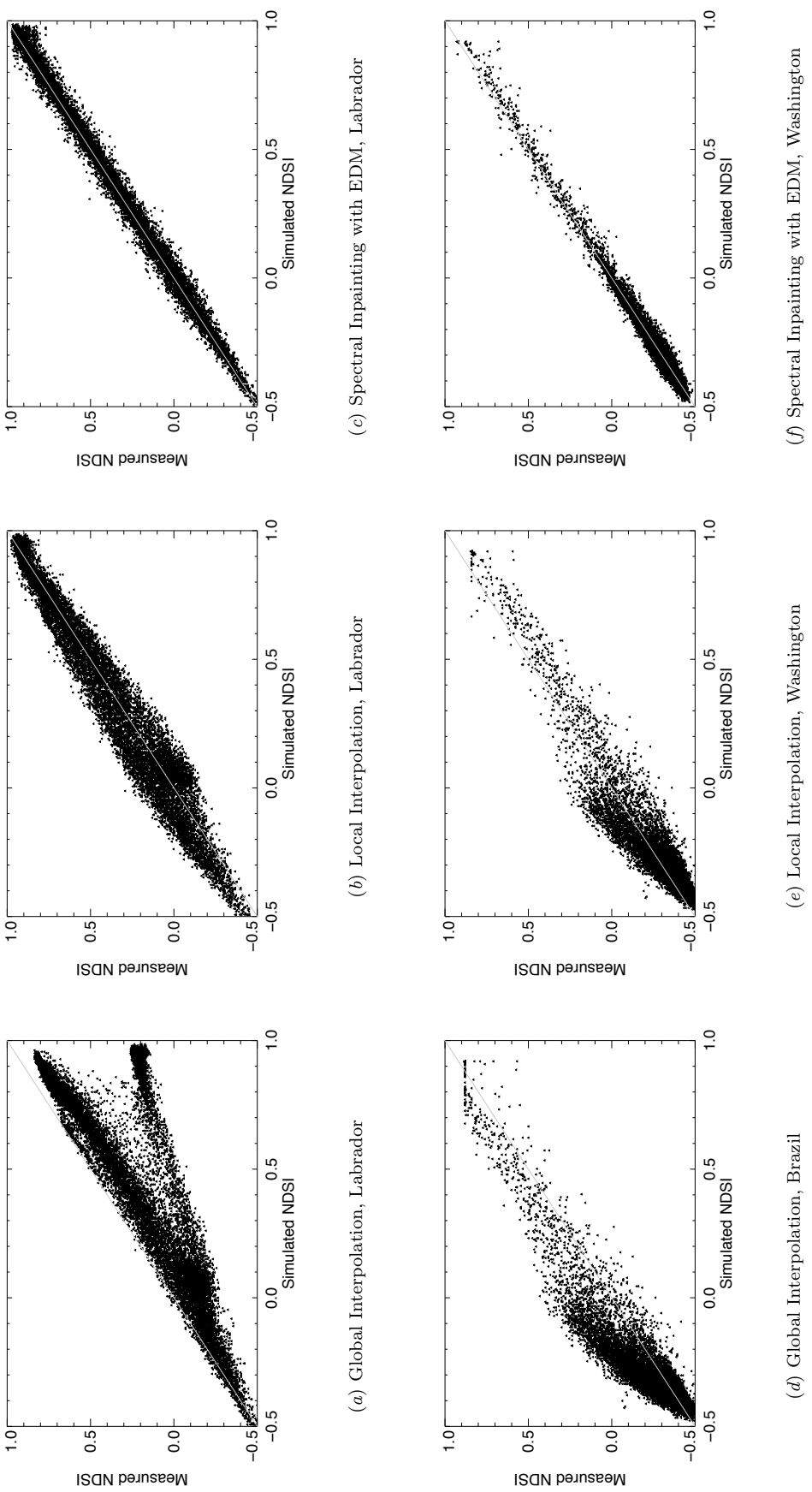
The gaussian weighting in the previous equation ensures that pixels with different spectral characteristics are given weak weights. A question that arises is the selection of the parameter ϵ in (11) and the decay regulating parameter h in (12). In pratice, an alternative approach consists in imposing the cardinality of the set S instead of selecting a value for ϵ. The set S can be restricted to S_N, a set composed only of the N^{th} most spectrally similar pixels to the missing pixel x. For values of $N \ll \text{card}(S)$, the gaussian average of equation (12) can be replaced by a simple arithmetic average:

$$\hat{\rho_M}(x) = \frac{1}{N} \int_{S_N(x)} \rho_M(y) dy \tag{13}$$

Let us provide further details on the implementation aspect of the spectral inpainting method. In order to reduce the computational cost of the proposed algorithm, the entire image domain Ω used for the search of spectrally similar pixels can be replaced by a subdomain of limited size. The later is not necesseraly square, nor is it required to be centered at the missing pixel, although this is a reasonable choice to increase the probability of finding suitable pixels for the restoration procedure. A possible approach is then to fragment the acquired signal into blocks of $L \times L$ pixels and to apply the spectral inpainting on individual subimages where the search window is the actual entire subimage domain. The value of L should be as high as possible and can be adjusted with respect to the available computational power and/or required processing time.

3. Experimental results

To validate the spectral inpainting algorithm, we used MODIS Level 1B TOA reflectances at 1km resolution. Given the importance of MODIS band 6 for snow applications, we selected 6 images used in the seminal work of (Salomonson and Appel 2004) and captured by Terra MODIS over snow covered regions located in Alaska, Labrador, Siberia, Kuparuk, South America and Washington. Because Terra MODIS band 6 is not corrupted with line dropouts due to non functional detectors, values from 4 detectors have been set to zero to emulate missing lines of Aqua MODIS band 6. Spectral Similarity Measures have been computed using



(a) Global Interpolation, Labrador (b) Local Interpolation, Labrador (c) Spectral Inpainting with EDM, Labrador

(d) Global Interpolation, Brazil (e) Local Interpolation, Washington (f) Spectral Inpainting with EDM, Washington

Figure 3. Scatter plot of measured and simulated NDSI index with three restoration techniques, global interpolation ,local interpolation, spectral inpainting

bands 3 (459-479 μm), 4 (545-565μm), 5 (1230-1250μm) and 7 (2105-2155μm). Not only these bands are spectrally distant from each other but they also cover a wide range of wavelenght, which provides a good discrimination between geophysical elements with different spectral properties. Experiments (not illustrated here) based on the use of bands 5 and 7 alone, also indicate satisfactory results due to the high correlation between these bands and band 6 over snow covered areas. The value of N in (13), is fixed to $N = 1$. This choice is physically consistent because it correponds to the case where a missing pixel is given the same value as the most spectrally similar pixel in the same spectral band.
Global interpolation (Wang *et al.* 2006), local interpolation (Rakwatin *et al.* 2009) and the proposed spectral inpanting technique are applied to Terra MODIS band 6 with synthetic non functional detectors. Restored results for the image acquired over Alaska are ilustrated in figure 2. The distribution of missing pixels along scan lines offers a pratical aspect to visually evaluate the restoration quality. In fact, systematic residual striping and its magnitude on the restored images is a clear qualitative indicator on the accuracy of missing pixels estimation. In addition, quantitative analysis is conducted using the NDSI defined as:

$$\mathrm{NDSI} = \frac{\rho_4 - \rho_6}{\rho_4 + \rho_6} \tag{14}$$

Some of the images used in this study have been initially corrected using the destriping algorithm introduced in (Bouali and Ladjal 2011). Nevertheless, we noticed that on MODIS reflective bands, including band 6, stripe noise is mostly visible over low-radiance areas such as ocean targets and its impact on snow covered regions is not substantial. Scatter plots of measured and simulated NDSI values obtained with different restoration methodologies for images over Alaska and are illustrated in figure 2.2. Measured and simulated NDSI are also quantitatively compared in terms of Absolute Mean Error, Root Mean Square Error (RMSE), and correlation coefficient. Results reported in tables 3. indicate higher accuracy in the recovery of band 6 obtained with the spectral inpainting especially when the EDM is used as the spectral similarity measure.

4. CONCLUSION

In this letter, a restoration procedure for the retrieval of missing data from multispectral imagery is described and applied to the particular issue of Aqua MODIS band 6 non functional detectors. In order to distinguish different geophysical cover types, the methodology incorporates spectral similarity measures into a nonlocal inpainting strategy. Experiments conducted on Terra MODIS band 6 and comparison between simulated and measured NDSI shows improved results compared to global and local interpolation techniques. Studies dedicated to snow mapping and based on the use of Aqua MODIS can highly benefit from the proposed approach. Finally, we point out that the core principle of the spectral inpainting algorithm can also be applied for the denoising of multispectral/hyperspectral remotely sensed data.

Table 2. Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Correlation (COR) between simulated and measured NDSI values using different restoration techniques

–		**Alaska**			**Labrador**			**Siberia**		
Restoration		MAE	RMSE	COR	MAE	RMSE	COR	MAE	RMSE	COR
Global interpolation		0.0147	0.0336	0.9960	0.0770	0.1650	0.9232	0.0015	0.0302	0.9979
Local interpolation		0.0133	0.0254	0.9975	0.0220	0.0449	0.9927	0.0069	0.0147	0.9994
Spectral Inpainting	SAM	0.0103	0.0248	0.9973	0.0111	0.0265	0.9974	0.0077	0.0173	0.9991
	SIDM	0.0105	0.0255	0.9971	0.0112	0.0267	0.9974	0.0078	0.0175	0.9991
	EDM	0.0057	0.0127	0.9993	0.0084	0.0187	0.9987	0.0039	0.0083	0.9998

–		**Kuparuk**			**South America**			**Washington**		
Restoration		MAE	RMSE	COR	MAE	RMSE	COR	MAE	RMSE	COR
Global interpolation		0.0235	0.0495	0.9942	0.0398	0.0709	0.9887	0.0223	0.0439	0.9651
Local interpolation		0.0191	0.0446	0.9951	0.0568	0.0987	0.9815	0.0158	0.0366	0.9754
Spectral Inpainting	SAM	0.0099	0.0220	0.9987	0.0139	0.0314	0.9970	0.0119	0.0279	0.9853
	SIDM	0.0102	0.0231	0.9985	0.0126	0.0286	0.9975	0.0108	0.0249	0.9882
	EDM	0.0060	0.0131	0.9995	0.0065	0.0139	0.9994	0.0060	0.0134	0.9965

References

Baccini, A., Friedl, M.A., Woodcock, C.E., Warbington, R. and Remer, L., 2004, Forest biomass estimation over regional scales using multisource data. *J. Geophys. Res. Lett.*, **31**, pp. L10501.1–L10501.4.

Bouali, M. and Ladjal, S., 2011, Toward Optimal Destriping of MODIS Data using a Unidirectional Variational Model. *IEEE Transactions on Geoscience and Remote Sensing*, **49**, pp. 2924–2935.

Buades, A., Coll, B. and Morel, J.M., 2005, A review of image denoising algorithms, with a new one. *Simul*, **4**, pp. 490–530.

Efros, A. and Leung, T., 1999, Texture Synthesis by Non-parametric Sampling. In *Proceedings of the International Conference on Computer Vision*, pp. 1033–1038.

Ignatov, A., Minnis, P., Loeb, N., Wielicki, B., Miller, W., Sun-Mack, S., Tanr, D., Remer, L., Laszlo, I. and Geier, E., 2005, Two MODIS Aerosol Products over Ocean on the Terra and Aqua CERES SSF Datasets. *Journal of the Atmospheric Sciences*, **62**, pp. 1008–1031.

Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J. and Goetz, A.F.H., 1993, The spectral image processing system (SIPS)–interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, **44**, pp. 145 – 163.

Meer, F.V.D., 2006, The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. *International journal of applied earth observation and geoinformation*, **8**, pp. 3–17.

Rakwatin, P., Takeuchi, W. and Yoshifumi, Y., 2009, Restoration of Aqua MODIS Band 6 Using Histogram Matching and Local Least Squares Fitting. *IEEE Transactions on Geoscience and Remote Sensing*, **47**, pp. 613–627.

Salomonson, V.V. and Appel, I., 2004, Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote Sensing of Environment*, **89**, pp. 351–360.

Wang, L.L., Qu, J.J., Siong, X., Xianjun, H., Yong, X. and Nianzeng, C., 2006, A new method for retrieving band 6 of Aqua MODIS. *IEEE Geosci. Remote Sens. Lett.*, **3**, pp. 267270.

Yaroslavsky, L.P. and Eden, M., 1996, In *Proceedings of the Fundamentals of Digital Optics, Birkhauser Boston.*