Applied Mathematical Sciences, Vol. 10, 2016, no. 49, 2443 - 2450
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2016.64148

Quality Assessment of Restored Satellite Data Based on Signal to Noise Ratio

Asmala Ahmad

Department of Industrial Computing
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Melaka, Malaysia

Shaun Quegan

Department of Applied Mathematics
School of Mathematics and Statistics
University of Sheffield
Sheffield, United Kingdom

Abstract

A practical concept of assessing the quality of restored data based on signal to noise ratio (SNR) is reported. The data come from remote sensing satellite and has undergone restoration process due to atmospheric haze effects. The restoration involves removing haze mean due to haze scattering and haze randomness due to haze spatial variability. The results shows that the SNR of restored data can be computed if the haze mean and haze randomness components are known.

Keywords: Haze, Remote Sensing, Signal to Noise ratio

1 Introduction

Atmospheric haze causes visibility to drop, therefore affecting data acquired using optical sensors on board remote sensing satellites [7], [10], [11]. Haze modifies the spectral and statistical properties of remote sensing data so causing problems to data users [4], [5], [6]. This issue is particularly true for optical system such as Landsat (USA), SPOT (France) and RazakSAT (Malaysia) [1], [2], [3].

Degradation of satellite data is caused by two key components, haze scattering and signal attenuation [10], which can be represented by a statistical model. In [8], the statistical model for hazy satellite data can be expressed as:

$$L_i(V) = \left(1 - \beta_i^{(1)}(V)\right)T_i + L_O + \beta_i^{(2)}(V)H_i \tag{1}$$

where $L_i(V)$, T_i, H_i, L_o, $\beta_i^{(1)}$ and $\beta_i^{(2)}$ are the hazy dataset, the signal component, the pure haze component, the radiance scattered by the atmosphere, the signal attenuation factor and the haze weighting in satellite band i, respectively. H_i can be expressed as:

$$H_i = \overline{H_i} + H_{i_v} \tag{2}$$

Where $\overline{H_i}$ is the haze mean, which is assumed to be uniform within the image or sub-region of the image, and H_{i_v} is a zero-mean random variable corresponding to haze randomness. Hence:

$$\mathrm{Var}\left(H_{i_v}\right) = \mathrm{Var}\left(H_i\right) \tag{3}$$

So Equation (1) can be written as:

$$L_i(V) = \left[1 - \beta_i^{(1)}(V)\right]T_i + L_O + \beta_i^{(2)}(V)\left[\overline{H_i} + H_{i_v}\right] \tag{4}$$

In order to remove the haze effects [4], [5], we need to remove both the weighted haze mean $\beta_i^{(2)}(V)\overline{H_i}$ and the varying component $\beta_i^{(2)}(V)H_{i_v}$ and deal with the signal attenuation factor $\beta_i^{(1)}(V)$.

From [8], the effects of $\beta_i^{(1)}(V)$ to data quality are not significant, so we will not consider their removal throughout the analysis. We normally do not have prior knowledge about $\beta_i^{(2)}(V)\overline{H_i}$ therefore we need to estimate it from the hazy data itself. If the estimate is $\widehat{\beta_i^{(2)}(V)\overline{H_i}}$, subtracting it from $L_i(V)$ yields:

$$\widehat{L_{i_Z}(V)} = L_i(V) - \widehat{\beta_i^{(2)}(V)\overline{H_i}} = \left[1 - \beta_i^{(1)}(V)\right]T_i + L_O + \beta_i^{(2)}(V)\left[\overline{H_i} + H_{i_v}\right] - \widehat{\beta_i^{(2)}(V)\overline{H_i}} \tag{5}$$

Equation (5) becomes:

$$\widehat{L_{i_Z}(V)} = \left[1 - \beta_i^{(1)}(V)\right]T_i + \left[\beta_i^{(2)}(V)\overline{H_i} - \widehat{\beta_i^{(2)}(V)\overline{H_i}}\right] + \beta_i^{(2)}(V)H_{i_v} + L_O \tag{6}$$

where $\left[\beta_i^{(2)}(V)\overline{H_i} - \widehat{\beta_i^{(2)}(V)\overline{H_i}}\right]$ is the error associated with the difference between the ideal and estimated weighted haze mean.

A common way to measure the accuracy of restored data is to compare its quality with uncorrupted data [12], [13], [14]. Visual analysis offers a fast and simple way to do this, but suffers from possible analyst bias. Hence we propose two quantitative approaches to assess the quality of restored data.

2 Signal to Noise Ratio

One measure of performance for single band data is the signal-to-noise ratio (SNR), which quantifies how severely data have been degraded by noise [9]. SNR is defined as the ratio between the squared ratio of signal amplitude and noise amplitude:

$$\mathrm{SNR} = \left(\frac{A_S}{A_N}\right)^2 \tag{7}$$

where P_S and A_S are signal power and amplitude respectively, and similarly for noise. SNR also can be measured on a decibel scale (dB):

$$\mathrm{SNR}(\mathrm{dB}) = 10\log_{10}(\mathrm{SNR}) = 10\log_{10}\left(\frac{P_S}{P_N}\right) = 20\log_{10}\left(\frac{A_S}{A_N}\right) \tag{8}$$

The expression for SNR and its estimates vary between: (a) original hazy data (with nonzero-mean noise), (b) hazy data after subtracting the haze mean and (c) restored data (after filtering).

From Equation (1), the SNR of hazy data with nonzero-mean haze noise can be expressed as:

$$\begin{aligned}
\mathrm{SNR} &= \frac{\left\langle \left\{\left[1-\beta_i^{(1)}(V)\right]T_i + L_O\right\}^2 \right\rangle}{\left\langle \beta_i^{(2)2}(V) H_i^2 \right\rangle} \\
&= \frac{\left\langle \left[1-\beta_i^{(1)}(V)\right]^2 T_i^2 + 2\left[1-\beta_i^{(1)}(V)\right] T_i L_O + L_O^2 \right\rangle}{\left\langle \beta_i^{(2)2}(V) H_i^2 \right\rangle} \\
&= \frac{\left[1-\beta_i^{(1)}(V)\right]^2 \left\langle T_i^2 \right\rangle + 2L_O\left[1-\beta_i^{(1)}(V)\right]\left\langle T_i \right\rangle + L_O^2}{\left\langle \beta_i^{(2)}(V)^2 \left(\overline{H}_i + H_{i_v}\right)^2 \right\rangle} \\
&= \frac{\left[1-\beta_i^{(1)}(V)\right]^2 \left\langle T_i^2 \right\rangle + 2L_O\left[1-\beta_i^{(1)}(V)\right]\left\langle T_i \right\rangle + L_O^2}{\beta_i^{(2)}(V)^2 \left[\overline{H_i}^2 + \mathrm{Var}\left(H_{i_v}\right)\right]}
\end{aligned} \tag{9}$$

since by assumption $\beta_i^{(1)}(V)$ and $\beta_i^{(2)}(V)$ are the same for all pixels in the scene. Note that here we assume $\left[1-\beta_i^{(1)}(V)\right]T_i$ from the hazy data to be the signal amplitude because the effects of $\left[1-\beta_i^{(1)}(V)\right]$ to data quality is negligible; this applies for all cases. Due to the discrete properties of the hazy data, the exact values are replaced by their estimates:

$$\widehat{SNR} = \frac{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\left[1-\beta_i^{(1)}(V)\right]T_i + L_O\right\}^2}{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\beta_i^{(2)}(V)^2\left(\overline{H}_i + H_{i_v}\right)^2} \tag{10}$$

where Q_m and Q_n are the numbers of pixels in the rows and columns of the image respectively. Note that such calculation is only possible if the values of T_i, $\overline{H}_i$, H_{i_v}, $\beta_i^{(1)}(V)$, $\beta_i^{(2)}(V)$, Q_m and Q_n are known apriori (e.g. simulated dataset). The exact SNR of degraded data after subtraction of the weighted haze mean can be expressed as:

$$SNR = \frac{\left\langle\left\{\left[1-\beta_i^{(1)}(V)\right]T_i + L_O\right\}^2\right\rangle}{\left\langle\left\{\left[\beta_i^{(2)}(V)\overline{H_i} - \widehat{\beta_i^{(2)}(V)\overline{H_i}}\right] + \beta_i^{(2)}(V)H_{i_v}\right\}^2\right\rangle} \tag{11}$$

and can be estimated by:

$$\widehat{SNR} = \frac{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\left[1-\beta_i^{(1)}(V)\right]T_i + L_O\right\}^2}{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\left[\beta_i^{(2)}(V)\overline{H_i} - \widehat{\beta_i^{(2)}(V)\overline{H_i}}\right] + \beta_i^{(2)}(V)H_{i_v}\right\}^2} \tag{12}$$

Subsequently, the degraded data undergo spatial filtering. From Equation (5.9), for linear filtering, the exact SNR of restored data can be expressed as:

$$\mathrm{SNR}=\frac{\left\langle\left\{\left[1-\beta_i^{(1)}(V)\right]T_i+L_O\right\}^2\right\rangle}{\left\langle\left[\hat{f}(V)-\left(1-\beta_i^{(1)}(V)\right)T_i+L_O\right]^2\right\rangle}$$

$$=\frac{\left\langle\left\{\left[1-\beta_i^{(1)}(V)\right]T_i+L_O\right\}^2\right\rangle}{\left\langle\left[\begin{array}{l}\left(1-\beta^{(1)}(V)\right)h_{\text{linear}}(T_i)+h_{\text{linear}}\left(\left[\beta_i^{(2)}(V)\overline{H_i}-\widehat{\beta_i^{(2)}(V)\overline{H_i}}\right]\right)+\\ \beta_i^{(2)}(V)h_{\text{linear}}\left(H_{i_v}\right)+L_O-\left(1-\beta_i^{(1)}(V)\right)T_i-L_O\end{array}\right]^2\right\rangle}$$

$$=\frac{\left\langle\left\{\left[1-\beta_i^{(1)}(V)\right]T_i+L_O\right\}^2\right\rangle}{\left\langle\left\{\begin{array}{l}\left(1-\beta_i^{(1)}(V)\right)\left[h_{\text{linear}}(T_i)-T_i\right]+h_{\text{linear}}\left(\left[\beta_i^{(2)}(V)\overline{H_i}-\widehat{\beta_i^{(2)}(V)\overline{H_i}}\right]\right)+\\ \beta_i^{(2)}(V)h_{\text{linear}}\left(H_{i_v}\right)\end{array}\right\}^2\right\rangle} \tag{13}$$

and can be estimated by:

$$\widehat{\mathrm{SNR}}=\frac{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\left[1-\beta_i^{(1)}(V)\right]T_i+L_O\right\}^2}{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\begin{array}{l}\left(1-\beta_i^{(1)}(V)\right)\left[h_{\text{linear}}(T_i)-T_i\right]+\\ h_{\text{linear}}\left(\left[\beta_i^{(2)}(V)\overline{H_i}-\widehat{\beta_i^{(2)}(V)\overline{H_i}}\right]\right)+\beta_i^{(2)}(V)h_{\text{linear}}(H_v)\end{array}\right\}^2} \tag{14}$$

For median filtering, the exact SNR can be expressed as:

$$\mathrm{SNR}=\frac{\left\langle\left\{\left[1-\beta_i^{(1)}(V)\right]T_i+L_O\right\}^2\right\rangle}{\left\langle\left[\hat{f}(V)-\left(1-\beta_i^{(1)}(V)\right)T_i\right]^2\right\rangle}$$

$$=\frac{\left\langle\left\{\left[1-\beta_i^{(1)}(V)\right]T_i+L_O\right\}^2\right\rangle}{\left\langle\left\{\begin{array}{l}\text{Median}\left(\begin{array}{l}\left[1-\beta_i^{(1)}(V)\right]T_i+\left[\beta_i^{(2)}(V)\overline{H_i}-\widehat{\beta_i^{(2)}(V)\overline{H_i}}\right]+\\ \beta_i^{(2)}(V)H_{i_v}+L_O\end{array}\right)-\\ \left[1-\beta_i^{(1)}(V)\right]T_i-L_O\end{array}\right\}^2\right\rangle} \tag{15}$$

and its estimate by:

$$\widehat{\text{SNR}} = \frac{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\left[1-\beta_i^{(1)}(V)\right]T_i + L_O\right\}^2}{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\begin{matrix}\text{Median}\left(\begin{matrix}\left[1-\beta_i^{(1)}(V)\right]T_i + \left[\beta_i^{(2)}(V)\overline{H_i} - \widehat{\beta_i^{(2)}(V)\overline{H_i}}\right] + \\ \beta_i^{(2)}(V)H_{i_v} + L_O\end{matrix}\right) - \\ \left[1-\beta_i^{(1)}(V)\right]T_i - L_O\end{matrix}\right\}^2} \quad (16)$$

3 The SNR of Restored Data when the Haze Mean is Known Exactly

When the haze mean is known exactly, $\left[\beta_i^{(2)}(V)\overline{H_i} - \widehat{\beta_i^{(2)}(V)\overline{H_i}}\right] = 0$ and therefore can be eliminated. Hence the SNR after subtraction of the haze mean is:

$$\widehat{\text{SNR}} = \frac{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\left[1-\beta_i^{(1)}(V)\right]T_i + L_O\right\}^2}{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\beta_i^{(2)}(V)^2 H_{i_v}{}^2} \quad (17)$$

For linear filtering we have:

$$\widehat{\text{SNR}} = \frac{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\left[1-\beta_i^{(1)}(V)\right]T_i + L_O\right\}^2}{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\left[1-\beta_i^{(1)}(V)\right]\left[h_{\text{linear}}(T_i) - T_i\right] + \beta_i^{(2)}(V)h_{\text{linear}}\left(H_{i_v}\right)\right\}^2} \quad (18)$$

For median filtering we have:

$$\widehat{\text{SNR}} = \frac{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\left[1-\beta_i^{(1)}(V)\right]T_i + L_O\right\}^2}{\sum_{m=1}^{Q_m}\sum_{n=1}^{Q_n}\left\{\begin{matrix}\text{Median}\left(\left[1-\beta_i^{(1)}(V)\right]T_i + \beta_i^{(2)}(V)H_{i_v} + L_O\right) - \\ \left[1-\beta_i^{(1)}(V)\right]T_i - L_O\end{matrix}\right\}^2} \quad (19)$$

4 Conclusion

In this paper, we have proposed a general concept of assessing the quality of restored data based on SNR. The SNR of restored data depends very much on the a priori knowledge of the haze mean and haze randomness components. These components increase as visibility decreases and therefore need to be known in order to remove haze and finally to estimate the SNR of restored data.

Acknowledgements. We would like to thank Universiti Teknikal Malaysia Melaka for funding this study under FRGS Grant (FRGS/2/2014/ICT02/FTMK/02/F00245) and Agency Remote Sensing Malaysia for providing the data.

References

[1] A. Ahmad, Classification Simulation of RazakSAT Satellite, *Procedia Engineering*, **53** (2013), 472 – 482.
http://dx.doi.org/10.1016/j.proeng.2013.02.061

[2] A. Ahmad and S. Quegan, Analysis of maximum likelihood classification technique on Landsat 5 TM satellite data of tropical land covers, *Proceedings of 2012 IEEE International Conference on Control System, Computing and Engineering (ICCSCE2012)*, (2012), 280 – 285.
http://dx.doi.org/10.1016/j.proeng.2013.02.061

[3] A. Ahmad and S. Quegan, Comparative analysis of supervised and unsupervised classification on multispectral data, *Applied Mathematical Sciences*, **7** (2013), no. 74, 3681 – 3694.
http://dx.doi.org/10.12988/ams.2013.34214

[4] A. Ahmad and Mohd Khanapi Abdul Ghani, Haze reduction in remotely sensed data, *Applied Mathematical Sciences*, **8** (2014), no. 36, 1755 – 1762.
http://dx.doi.org/10.12988/ams.2014.4289

[5] A. Ahmad and S. Quegan, The Effects of haze on the spectral and statistical properties of land cover classification, *Applied Mathematical Sciences*, **8** (2014), no. 180, 9001 – 9013.
http://dx.doi.org/10.12988/ams.2014.411939

[6] A. Ahmad and S. Quegan, The effects of haze on the accuracy of satellite land cover classification, *Applied Mathematical Sciences*, **9** (2015), no. 49, 2433 – 2443. http://dx.doi.org/10.12988/ams.2015.52157

[7] A. Asmala, M. Hashim, M. N. Hashim, M. N. Ayof and A. S. Budi, The use of remote sensing and GIS to estimate Air Quality Index (AQI) Over Peninsular Malaysia, GIS Development, (2006), 5.

[8] A. Ahmad and S. Quegan, Haze modelling and simulation in remote sensing satellite data, *Applied Mathematical Sciences*, **8** (2014), no. 159, 7909 – 7921. http://dx.doi.org/10.12988/ams.2014.49761

[9] J. R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*, Pearson Prentice Hall, New Jersey, USA, 1996.

[10] M. F. Razali, A. Ahmad, O. Mohd and H. Sakidin, Quantifying haze from satellite using haze optimized transformation (HOT), *Applied Mathematical Sciences*, **9** (2015), no. 29, 1407 – 1416. http://dx.doi.org/10.12988/ams.2015.5130

[11] M. Hashim, K. D. Kanniah, A. Ahmad, A. W. Rasib, Remote sensing of tropospheric pollutants originating from 1997 forest fire in Southeast Asia, *Asian Journal of Geoinformatics*, **4** (2004), 57 – 68.

[12] M. Story and R. Congalton, Accuracy assessment: a user's perspective, *Photogrammetric Engineering and Remote Sensing*, **52** (1986), 397 – 399.

[13] U. K. M. Hashim and A. Ahmad, The effects of training set size on the accuracy of maximum likelihood, neural network and support vector machine classification, *Science International-Lahore*, **26** (2014), no. 4, 1477 – 1481.

[14] J. R. Thomlinson, P. V. Bolstad, and W. B. Cohen, Coordinating methodologies for scaling landcover classifications from site-specific to global: steps toward validating global map products, *Remote Sensing of Environment*, **70** (1999), 16 – 28. http://dx.doi.org/10.1016/s0034-4257(99)00055-3

Received: April 28, 2016; Published: July 28, 2016