In: Wagner W., Székely, B. (eds.): ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, July 5–7, 2010, IAPRS, Vol. XXXVIII, Part 7A

FOUR REDUCED-REFERENCE METRICS FOR MEASURING HYPERSPECTRAL IMAGES AFTER SPATIAL RESOLUTION ENHANCEMENT

S. Qian [a], *, G. Chen [a]

[a] Canadian Space Agency, 6767 route de l'Aeroport, St-Hubert, Quebec, Canada J3Y8Y9
(shen-en.qian, guangyi.chen)@asc-csa.gc.ca

KEY WORDS: Hyper spectral, Image, Spatial, Resolution, Metric

ABSTRACT:

In this paper, four new reduced-references (RR) metrics are proposed for measuring the visual quality of hyperspectral images after having undergone spatial resolution enhancement. These metrics can measure the visual quality of hyperspectral images whose full-reference (FR) image is not available whereas the low spatial resolution reference image is available. A FR metric requires the reference image and the test image to have the same size. After spatial resolution enhancement of hyperspectral images, the size of the enhanced images is larger than that of the original image. Thus, the FR metric cannot be used. A common approach in practice is to first down-sample an original image to a low resolution image, then to spatially enhance the down-sampled low resolution image using an enhancement technique. In this way, the original image and the enhanced image have the same size and the FR metric can be applied to them. However, this common approach can never directly assess the image quality of the spatially enhanced image that is produced directly from the original image. Experimental results showed that the proposed RR metrics work well for measuring the visual quality of spatial resolution enhanced hyperspectral images. They are consistent with the corresponding FR metrics.

1. INTRODUCTION

Measurement of image quality is of fundamental importance to many image processing applications. Image quality assessment algorithms are in general classified into three categories: full-reference (FR), reduced-reference (RR), and no-reference (NR) algorithms. True NR algorithms are very difficult to design and little progress has been made (Sheikh et al, 2005). FR algorithms are easier to design and the majority of image quality assessment algorithms are of this type. In FR quality assessment, a reference image of perfect quality is assumed to be available. However, in RR or NR quality assessment, partial or no reference information is available.

Mean square error (MSE) is the simplest FR metric between the reference image x and the processed image y:

$$MSE(x,y)=\frac{1}{N}\sum_{i,j}(y(i,j)-x(i,j))^2 \tag{1}$$

where N is the total number of pixels in the images x and y. The MSE is easy to compute and implement in software and hardware. However, the MSE is not a good image quality measure as it is not well matched to perceived image quality. Two distorted images with the same MSE may have very different types of errors, some of which are more visible than others. Thus one image may look very much similar to the reference, whereas another may look very much distorted.

Peak signal to noise ratio (PSNR) is also a popular FR metric to measure the quality of a reconstructed image, and it is defined as:

$$PSNR(x,y)=10\log_{10}\left(\frac{\max(x)^2}{\frac{1}{N}\sum_{i,j}(y(i,j)-x(i,j))^2}\right) \tag{2}$$

The PSNR has been used as a standard metric in image denoising and other related image processing tasks.

Wang and Bovik (2002) proposed the Q index for a reference image x and an image y to be evaluated,

$$Q(x,y)=\frac{\sigma_{xy}}{\sigma_x\sigma_y}\cdot\frac{2\mu_x\mu_y}{\mu_x^2+\mu_y^2}\cdot\frac{2\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2} \tag{3}$$

where μ_x and μ_y are sample means, σ_x^2 and σ_y^2 are sample variances, and σ_{xy} is the sample cross-covariance between x and y. The Q index is a FR metric and it is easy to calculate and applicable to various image processing applications. It outperforms the MSE significantly under different types of image distortions. Wang et al (2004) also developed the structural similarity (SSIM) index, which is also a FR metric, by comparing local correlations in luminance, contrast, and structure between the reference and distorted images. The SSIM index is defined as:

$$SSIM(x,y)=\frac{\sigma_{xy}+C_1}{\sigma_x\sigma_y+C_1}\cdot\frac{2\mu_x\mu_y+C_2}{\mu_x^2+\mu_y^2+C_2}\cdot\frac{2\sigma_x\sigma_y+C_3}{\sigma_x^2+\sigma_y^2+C_3} \tag{4}$$

where μ_x and μ_y are sample means of images x and y, σ_x^2 and σ_x^2 are sample variances, and σ_{xy} is the sample cross-covariance between x and y. The constants C_1, C_2, C_3 stabilize SSIM when the means and variances become small. The mean SSIM (MSSIM) over the whole image gives the final quality measure.

* Corresponding author

Sheikh and Bovik (2006) developed a visual information fidelity (VIF) index for FR measurement of image visual quality. Let $\boldsymbol{e}=\boldsymbol{c}+\boldsymbol{n}$ be the reference image**,** and $\boldsymbol{n}$ zero-mean normal distribution $N(0,\sigma_n^2 I)$ noise. Also, let $\boldsymbol{f}=\boldsymbol{d}+\boldsymbol{n}'= \boldsymbol{g}\boldsymbol{c}+\boldsymbol{v}'+\boldsymbol{n}'$ be the test image, where $\boldsymbol{g}$ represents the blur, $\boldsymbol{v}'$ the additive zero-mean Gaussian white noise with covariance $\sigma_v^2 I$, and $\boldsymbol{n}'$ the zero-mean normal distribution $N(0,\sigma_n^2 I)$ noise. Then, VIF can be computed as the ratio of the mutual information between $\boldsymbol{c}$ and $\boldsymbol{f}$, and the mutual information between $\boldsymbol{c}$ and $\boldsymbol{e}$ for all wavelet subbands except the lowest approximation subband.

$$VIF = \frac{\sum I(c;f \mid z)}{\sum I(c;e \mid z)} \tag{5}$$

All the metrics above are popular metrics published in the literature for FR image quality assessment. However, they require the reference image and the test image to have the same image size. After spatial resolution enhancement of hyperspectral images, the size of the enhanced images is larger than that of the original image. Thus, these metrics cannot be used to assess the quality of the enhanced images. A common approach in practice is to first down-sample an original image to a low resolution image, then to spatially enhance the down-sampled low resolution image using an enhancement technique. In this way, the original image and the enhanced image have the same size and the FR metrics can be applied to them. However, this common approach can never directly assess the image quality of the spatially enhanced image that is produced directly from the original image. The image quality of the enhanced image measured based on the down-sampled low resolution image may or may not reflect the real quality of the image that is enhanced directly from the original image, as the down-sampling procedure introduces artificial effects.

This paper proposes new RR metrics. A brief review about the RR metric is given here. Wang and Simoncelli (2005) proposed an RR image quality assessment method based on a natural image statistic model in the wavelet transform domain. They used the Kullback-Leibler distance between the marginal probability distributions of wavelet coefficients of the reference and distorted images as a measure of image distortion. A generalized Gaussian model was employed to summarize the marginal distribution of wavelet coefficients of the reference image, so that only a relatively small number of RR features are needed for the evaluation of image quality. Li and Wang (2009) proposed an RR algorithm using statistical features extracted from a divisive normalization-based image representation. They demonstrated that such an image representation has simultaneous perceptual and statistical relevance and its statistical properties are significantly changed under different types of image distortions. Engelkea et al (2009) developed RR objective perceptual image quality metrics for use in wireless imaging. Instead of focusing only on artifacts due to source encoding, they followed an end-to-end quality approach that accounts for the complex nature of artifacts that may be induced by a wireless communication system.

In this paper, four new RR metrics were proposed for measuring the image fidelity of a testing image that has higher spatial resolution (i.e. larger size than that of the original image). It is assumed that a low spatial resolution reference image is available, whereas the high spatial resolution reference image is not. These four proposed RR metrics do not require the sizes of the reference image and the test image to be the same.

The iterative back projection (IBP) (Irani and Peleg, 1991, 1993) technique was chosen to enhance the spatial resolution of testing hyperspectral images in order to demonstrate the usefulness of these metrics. Experimental results reported in section 3 show that the proposed metrics can measure the image quality of the spatial resolution enhanced images very well.

2. CONSTRUCTING NEW RR METRICS FROM EXISTING FR METRICS

In this section, four new RR metrics are proposed for assessing the image quality of a spatial resolution enhanced image. They can be derived as follows.

Let the size of the low spatial resolution image $\boldsymbol{f}$ be $P{\times}Q$, and the size of the corresponding spatial resolution enhanced image $\boldsymbol{g}$ be $2P{\times}2Q$. This means that the spatial resolution of image $\boldsymbol{f}$ is enhanced at a factor of $2{\times}2$. The following four down-sampled images at a factor of $2{\times}2$, can be defined as:

$$g_{11} = g(1:2:2P, 1:2:2Q) \tag{6}$$
$$g_{12} = g(1:2:2P, 2:2:2Q) \tag{7}$$
$$g_{21} = g(2:2:2P, 1:2:2Q) \tag{8}$$
$$g_{22} = g(2:2:2P, 2:2:2Q) \tag{9}$$

where $g(i:2:2P, j:2:2Q)$, ($i=1,2;\ j=1,2$), is a matrix which starts at the pixel (i,j) of image $\boldsymbol{g}$ and extract every other pixels in $\boldsymbol{g}$ along both the x and the y directions with a step of 2. Since the low spatial resolution image $\boldsymbol{f}$ and the images $g_{i,j}$ ($i,j=1,2$) have the same image size, one can use any FR metrics to measure the image quality between them. The following four RR metrics are proposed in this paper:

$$PSNR(f;g) = \frac{1}{4}\sum_{i=1}^{2}\sum_{j=1}^{2} PSNR(f;g_{ij}) \tag{10}$$
$$Q(f;g) = \frac{1}{4}\sum_{i=1}^{2}\sum_{j=1}^{2} Q(f;g_{ij}) \tag{11}$$
$$MSSIM(f;g) = \frac{1}{4}\sum_{i=1}^{2}\sum_{j=1}^{2} MSSIM(f;g_{ij}) \tag{12}$$
$$VIF(f;g) = \frac{1}{4}\sum_{i=1}^{2}\sum_{j=1}^{2} VIF(f;g_{ij}) \tag{13}$$

Experimental results conducted in the next section show that these four RR metrics can measure the image quality of a spatial resolution enhanced image very well. Even though these four RR metrics are derived for a special spatial resolution enhancement factor 2×2, it is easy to extend it to other spatial resolution enhancement factor $M{\times}N$, where both M and N are positive integers.

The IBP is chosen to enhance the spatial resolution of the testing images. For simplicity, this paper only considers spatially increasing the resolution by a factor of 2×2. It is easy to extend IBP to even higher resolution enhancement. IBP consists of two steps: (i) projection, and (ii) back-projection. It enhances spatial resolution of an image by performing projection and back-projection iteratively until satisfactory results are obtained. In IBP, the imaging is regarded as a projecting process that includes shifting, under-sampling and blurring operations to generate a set of low resolution images. So the reconstruction of a high-resolution image from these low resolution images can then be regarded as a back-projecting process which includes de-blurring, up-sampling and anti-shifting operations. This back-projection is performed in an iterative way. The IBP algorithm converges rapidly, and can meet the need of real-time processing since it only deals with some simple operations. Generally, the resultant image has satisfactory visual effect after 10 iterations.

For the sake of comparison with IBP, interpolation is used to enhance the spatial resolution of the testing images. The bilinear interpolation was chosen as the interpolation method in the experiments.

3. EXPERIMENTAL RESULTS

In this section, a number of experiments were conducted to demonstrate the feasibility of the proposed RR metrics. Three hyperspectral data cubes were tested in this paper. The 2-dimensional (2D) band images of the data cubes are used to test the proposed RR metrics. The first hyperspectral data cube was acquired using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) in the Cuprite mining district, Nevada, in 1997. The original scene with size of 614×512 pixels and 224 bands is available online at http://aviris.jpl.nasa.gov/html/aviris.freedata.html. The upper-right corner of the scene that consists of 350×350 pixels and 224 bands are cut out. This scene is well understood mineralogically and it has been made a standard test site for validation and assessment of remote sensing methods (Chen and Qian, 2007, 2008a, 2008b, 2009; Wang and Chang, 2006). Due to water absorption and low SNR, the bands 1-3, 105-115 and 150-170 are removed. As a result, a total of 189 bands are used in our experiments. Figure 1 shows the image of the Cuprite data cube at wavelength 827 nm (band #50).

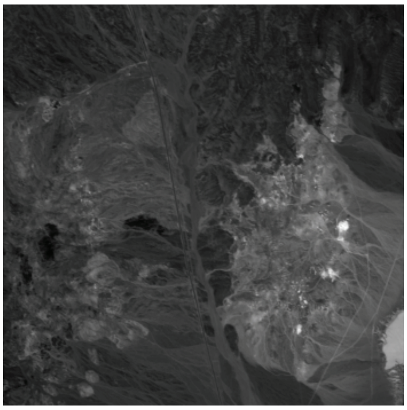

Figure 1. The AVIRIS Cuprite scene at displayed at wavelength 827nm (spectral band #50).

The second hyperspectral data cube was acquired using the airborne Short-wave-infrared Full Spectrum Imager II (SFSI-II). The data cube was collected over Key Lake in northern Saskatchewan, Canada for studying the capability of imaging spectrometers in identifying uranium mine and associated activities. The data cube was acquired with a ground sample distance (GSD) of 3.19m×3.13m. The size of the data cube is 1090 lines by 496 pixels by 240 bands. The scene of the testing data cube includes a mill complex and a mine complex. Figure 2 shows an image at wavelength 1304 nm (band #16) of this data cube.

The third hyperspectral data cube was also collected using the SFSI-II for studying target detection from short wave infrared hyperspectral imagery. The GSD of the data cube is 2.20m×1.85m. The size of the data cube is 140 lines by 496 pixels by 240 bands. Man-made targets with different materials and sizes were deployed in a mixed of sand and low-density grass cover within the scene of the data cube. Seven pieces of awnings with varying sizes ranging from 12m×12m to 0.2m×0.2m, four pieces of polythene, four pieces of white tarp and four pieces of white cotton with varying size ranging from 6m×6m to 0.5m×0.5m were deployed. In addition, a 3m×3m piece of white tarp was placed on a large vinyl turf mat of size 11m×14m. Figure 3 shows a region-of-interest (size: 140×140) at wavelength 1289 nm (band #13) of this data cube.

Figure 2. The Key Lake scene displayed at wavelength 1304 nm (spectral band #16).

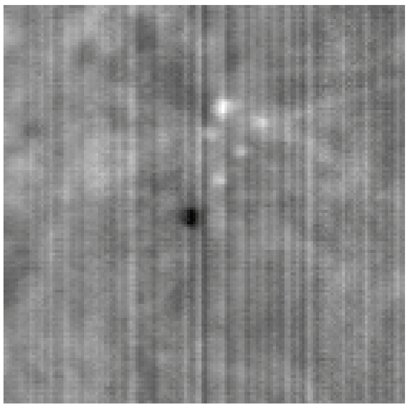

Figure 3. The SFSI-II data cube with man-made targets displayed at wavelength 1289 nm (spectral band #13).

Ideally, the IBP and bilinear interpolation should be used to enhance the spatial resolution of every spectral band in the hyperspectral data cubes, and then the proposed four RR metrics were used to assess their image quality. However, due to the workload of computing the IBP and the metrics for the whole datacubes, as a consequence, bands #16, #50 and #13 have been chosen in the experiments for the Cuprite data cube, the Key Lake data cube, and the Target data cube, respectively.

The PSNR, Q index, MSSIM, and VIF are of widely used FR metrics in image processing. In the experiments, these four FR metrics are compared to their corresponding RR metrics. Table 4 lists the experimental results of the metrics applied to the spatial resolution enhanced images by using IBP and interpolation. The IBP was run for 30 iterations in order to generate a higher quality spatial resolution enhanced image.

Data cube	Spatial Enhancement Method	Full-Reference metrics				Reduced-Reference metrics			
		PSNR	Q	MSSIM	VIF	PSNR	Q	MSSIM	VIF
Cuprite	IBP	36.51	0.82	0.91	0.69	43.82	0.97	0.99	0.87
	Interpolation	35.67	0.78	0.90	0.48	38.37	0.92	0.96	0.75
Key Lake	IBP	34.87	0.75	0.89	0.77	40.11	0.96	0.98	0.87
	Interpolation	32.41	0.70	0.87	0.54	34.59	0.87	0.94	0.78
Target datacube	IBP	53.33	0.78	0.99	0.77	61.67	0.97	1.00	0.98
	Interpolation	53.14	0.74	0.99	0.65	56.35	0.89	1.00	0.88

Table 4. Experimental results of four FR image quality metrics and the four proposed RR metrics of the test images that are spatially enhanced by using the IBP and interpolation methods. For the FR metrics, a test image is first down-sampled at a factor of 2×2, then spatially enhanced at a factor of 2×2. For the proposed RR metrics, a test image is spatially enhanced by a factor of 2×2 without a prior down-sampling.

For the FR metrics, a test image is first down-sampled at a factor of 2×2, then is spatially enhanced at a factor of 2×2 in order to satisfy the requirement of the processed image having the same size as the reference image. For the proposed RR metrics, an original test image is spatially enhanced at a factor of 2×2 without a prior down-sampling. From the table, it can be seen that IBP-based method always produces better results than the bilinear interpolation no matter whether the original image is down-sampled or not. More importantly, the proposed RR metrics measure the image quality of the spatial resolution enhanced images very well, and they are consistent with the corresponding FR metrics. This indicates that the proposed RR metrics are reliable metrics for measuring the quality of the spatial resolution enhanced images.

4. CONCLUSION

In this paper, four new RR metrics are proposed to measure the quality of the spatially enhanced hyperspectral images. These metrics do not require the sizes of the reference and test images to be the same. However, all FR metrics published in the literature require both images to have the same size. The IBP and bilinear interpolation are used to increase the spatial resolution of a testing image. Experimental results show that the proposed four RR metrics can measure the image quality of the spatial resolution enhanced images very well. Even though only hyperspectral images are tested in this paper, the proposed metrics can be used to measure the image quality of any other spatially enhanced images.

REFERENCES

Chen, G. Y. and Qian, S. E., 2007. Dimensionality reduction of hyperspectral imagery using improved locally linear embedding. *Journal of Applied Remote Sensing*, 1, pp. 013509.

Chen, G. Y. and Qian, S. E., 2008a. Evaluation and comparison of dimensionality reduction methods and band selection. *Canadian Journal of Remote Sensing,* 34(1), pp. 26-32.

Chen, G. Y. and Qian, S. E., 2008b. Simultaneous dimensionality reduction and denoising of hyperspectral imagery using bivariate wavelet shrinking and principal component analysis. *Canadian Journal of Remote Sensing*, 34(5), pp. 447-454.

Chen, G. Y. and Qian, S. E., 2009. Denoising and dimensionality reduction of hyperspectral imagery using wavelet packets, neighbour shrinking and principal component analysis. *International Journal of Remote Sensing*, 30(18), pp. 4889-4895.

Engelkea, U., Kusuma, M., Zepernick, H. J. and Caldera, M., 2009. Reduced-reference metric design for objective perceptual quality assessment in wireless imaging. *Signal Processing: Image Communication*, 24(7), pp. 525-547.

Irani, M. and Peleg, S., 1991. Improving resolution by image registration. *Computer Vision, Graphics and Image Processing*, 53, pp. 231-239.

Irani, M. and Peleg, S., 1993. Motion analysis for image enhancement: resolution, occlusion and transparency. *Journal of Visual Communications and Image Representation*, 4, pp. 324-335.

Li, Q. and Wang, Z., 2009. Reduced-reference image quality assessment using divisive normalization-based image representation. *IEEE Journal of Selected Topics in Signal Processing*, *Special issue on Visual Media Quality Assessment*, 3(2), pp. 202-211.

Sheikh, H. R., Bovik, A. C. and Cormack, L., 2005. No-reference quality assessment using natural scene statistics: JPEG2000. *IEEE Transactions on Image Processing*, 14(11), pp. 1918-1927.

Sheikh, H. R. and Bovik, A. C., 2006. Image information and visual quality. *IEEE Transactions on Image Processing*, 15(2), pp. 430-444.

Wang, Z. and Bovik, A. C., 2002. A universal image quality index. *IEEE Signal Processing Letters*, 9, pp. 81-84.

Wang, Z., Bovik, A. C., Sheikh, H. R. and Simoncelli, E. P., 2004. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), pp. 600-612.

Wang, J. and Chang, C.-I., 2006. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6), pp. 1586-1600.

Wang, Z. and Simoncelli, E. P., 2005. Reduced-reference image quality assessment using a wavelet-domain natural image statistic model. *Human Vision and Electronic Imaging X, Proc. SPIE*, vol. 5666, San Jose, CA.

ACKNOWLEDGEMENTS

This work was partially supported by the Visiting Fellowship in Canadian Government Laboratory coordinated by the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors would like to thank J. Levesque at Defence Research and Development Canada for providing the test datacubes.