QUALITY ASSESSMENT OF PAN-SHARPENING METHODS

Gintautas Palubinskas

German Aerospace Center DLR, Remote Sensing Technology Institute
Oberpfaffenhofen, 82234 Wessling, Germany

ABSTRACT

The quality of pan-sharpened image is usually quantified separately by various spectral and spatial quality measures mostly originating from image processing. This quantity and diversity of quality measures makes it quite difficult to rank different image fusion methods. A new Joint Quality Measure (JQM) is proposed which is based on the combination of two Structural SIMilarity (SSIM) indices (one for spectral quality and another one for spatial quality) allowing comparison of different methods using a sole measure. Quality assessment of four fusion methods: Component Substitution (CS), General Fusion Filtering (GFF), variant of GFF and ATWT (one of ARSIS implementations) is performed for IKONOS and WorldView-2 satellite optical remote sensing data. Experiments and results showed the superiority of the proposed JQM when compared with already known joint measure - Quality with No Reference (QNR) index. Moreover, the results are fully supported by a visual analysis of imagery.

Index Terms— Multi-resolution, multi-sensor, image fusion, pan-sharpening, quality measure

1. INTRODUCTION

Multi-resolution image fusion also known as pan-sharpening aims to include spatial information from a high resolution image, e.g. panchromatic or Synthetic Aperture Radar (SAR) image, into a low resolution image, e.g. multi-spectral or hyper-spectral image, while preserving spectral properties of a low resolution image. A large number of algorithms and methods to solve this problem were introduced during the last two decades which can be divided into two large groups. First group of methods is based on a linear spectral transformation (e.g. Intensity-Hue-Saturation, Principal Component Analysis, Gram-Schmidt ortho-normalization) followed by a Component Substitution (CS) [1]. Methods of the second group use spatial frequency decomposition usually performed by means of high pass filtering e.g. boxcar filter in signal domain [2], filtering in Fourier domain [3,4] or multi-resolution analysis using wavelet transform [5,6]. Here we have to mention that there are some attempts to combine both types of methods [7].

In parallel to pan-sharpening methods development many attempts were undertaken to assess quantitatively their quality usually using measures originating from image processing such as Mean Squared Error (MSE), Correlation Coefficient (CC), Structural SIMilarity (SSIM) index [8] and finally recently proposed joint measures: product of spectral measure based on SSIM and spatial measure based on CC [9] and Quality with No Reference (QNR) measure [10]. Moreover, statistical evaluation of most popular pan-sharpening quality assessment measures was performed in [11,12] showing most relevant and practical quality measures using separate spectral and spatial image quality assessment over a broad collection of optical satellite sensor data. These results served as an inspiration for a new joint quality assessment measure (an enhanced version of a measure already presented in [4]) which is the main topic of this paper. This new sole quality measure allows assessment of a general image quality and provides an easy and practical way to rank various fusion methods.

2. METHODS

The following four pan-sharpening methods: General Fusion Filtering (GFF) [3], variant of General Fusion Filtering (HPFM) [4], Component Substitution (CS) [1] and implementation of ARSIS concept - ATWT Model for Wavelet Fusion [5,6,13] - are investigated in this paper. First two methods are based on filtering operation and thus are very well suited for producing different quality imagery, useful for investigating various quality assessment measures.

3. QUALITY MEASURES

The quality of pan-sharpening is usually measured by spectral and/or spatial quality measures to cover both attributes of a processing result.

3.1. Spectral quality measure

Recent comparison [11,12] for various optical satellite data showed that the correlation CC or SSIM between original spectral bands and corresponding low pass filtered pan-sharpened bands is quite well suited for this purpose. It

Proc. IGARSS, 13-18 July, 2014, Québec, Canada, IEEE

allows us to measure a spectral quality or preservation of a pan-sharpening method for individual bands or by averaging for all bands. It has high values (optimal value is 1) for a good spectral preservation and low (or negative) values for low spectral characteristics preservation. We have to note that SSIM could be more preferable for data of different sensors, thus a more general measure which will be called Quality for Low Resolution (QLR) is used

$$QLR = \frac{1}{n}\sum_{k=1}^{n} SSIM(ms_k, (fms_k * lpf)\downarrow), \tag{1}$$

where (high resolution data)$\downarrow$ means subsampling of high resolution data to low resolution data, *ms* – low resolution data, *fms* – fused data, *lpf* – low pass filter, *n* – number of bands and *k* – band index.

This measure alone is not able to assess the quality of fusion result, because it is calculated only in a reduced (low) image resolution/scale.

3.2. Spatial quality measure

The same investigation [11,12] showed a preference of the structural similarity index measure SSIM between original panchromatic band (*pan*) and pan-sharpened bands (*fms*) for a spatial quality assessment. It exhibits high values (optimal value 1) for a high spatial quality and low values for low spatial quality. Here we have to note that due to different spectra of multispectral and panchromatic data the correlation as proposed by [9] may be not sufficient because of possible mean and standard deviation differences. SSIM allows us to account for such differences much better. Moreover, under the assumption that interpolated spectral bands *msi* and *pan* are highly correlated, a new spatial quality measure can be introduced: SSIM between intensity image calculated from fused multispectral bands and original panchromatic band, which will be called Quality for High Resolution (QHR)

$$QHR = SSIM(pan, \frac{1}{n}\sum_{k=1}^{n} fms_k). \tag{2}$$

3.3. New joint quality measure (JQM)

In ideal case pan-sharpening method should exhibit both high spectral and spatial quality measure values. But it is not possible practically, because for example CS methods exhibit better spatial quality whereas filtering based methods deliver better spectral quality. Moreover, for GFF method (also valid for other filtering methods) different parameters (amount of filtering) lead to different image qualities. Thus a larger low pass filtering parameter value will lead to a higher spectral quality at the same time reducing spatial quality and vice versa. None of the known separate quality measures can fulfill above mentioned requirement as a sole measure. Thus, a joint quality measure could be helpful to achieve optimal parameter selection or best trade-off between spectral and spatial quality or find the best method for a particular application.

In this paper we propose a new Joint Quality Measure (JQM) which is based on QLR for spectral quality (1) and QHR for spatial quality (2). Due to different ranges of two measures we propose here to scale [4] one of the measures before averaging (producing a joint measure)

$$QHR_{norm} = \frac{QHR - QHR_{\min}}{QHR_{\max} - QHR_{\min}} \cdot (QLR_{\max} - QLR_{\min}) + QLR_{\min} \tag{3}$$

where QHR_{min} stands for minimum of all QHR values. For example, these values can be calculated for different filtering parameters of GFF method. Similarly other minimum and maximum values are defined. Analysis for different data, different methods (GFF and HPFM) and finally existing experience resulted in these typical extreme value ranges: 0.9-0.999 for QLR and 0.7-0.95 for QHR, which are used further to calculate JQM. Now the averaging of corresponding spectral and scaled spatial measures

$$JQM = (QLR + QHR_{norm})/2 \tag{4}$$

delivers the joint quality measure which is suitable for parameter selection and comparison of different pan-sharpening methods.

A new JQM is compared with a known joint measure - Quality with No Reference (QNR) [10].

4. EXPERIMENTS AND RESULTS

We shall illustrate our ideas concerning pan-sharpening quality assessment for two optical remote sensing sensors: IKONOS and WorldView-2 over two cities: Munich in South Germany and Melbourne in Australia. In this section we shall compare different methods: GFF, HPFM, CS and ATWT (see sect. 2) using the proposed JQM and already known QNR joint quality measures for different optical remote sensing data. JQM and QNR quality measure values for various methods and various parameter settings for IKONOS Munich data are presented in Table 1.

First, only interpolated multispectral data are evaluated. We see that cubic convolution (CC) interpolation method exhibits best (largest) spectral quality (QLR), what results in best (largest) JQM value. Bilinear (BIL) method is quite similar to CC. Nearest neighbor (NN) method produces best (largest) spatial quality (QHR), but due lower QLR value the JQM is lower than that of CC. Zero padding (ZP) method due its strong filtering properties (avoidance of aliasing) results in poorest quality values. This ranking corresponds quite well with visual analysis and existing experience. Now, looking at QNR ant its compound parts we can derive following conclusions. NN exhibits best (smallest) spectral quality (D_λ), followed quite closely by BIL and CC. ZP has worst spectral quality. These observations correspond quite well to QLR ranking except

NN. NN produces the best (smallest) spatial quality (D_s) too, thus resulting in best (largest) QNR value. This observation contradicts visual analysis and existing experience. Such ranking of quality cannot be accepted, and this is the **first argument** against this measure.

For filter based fusion methods (GFF and HPFM) JQM selects filtering parameters, which correspond well with visual analysis and existing experience. QNR prefers too high filtering parameters, thus selecting as best the low spatial quality images. QNR is unable to select correctly filtering parameter. This is the **second argument** against it. Further analysis of JQM results in following conclusions. For HPFM model type has no significant influence on the quality. Thus multiplicative model is excluded from further analysis. Moreover, usage of BIL and CC interpolation results in best quality, even better than GFF. This can be explained with the fact, that GFF uses ZP interpolation, which exhibits lower spatial quality in comparison to BIL or CC.

For CS, similarly as for HPFM, model type has no significant influence on the quality. In general, the JQM quality of CS is lower than that of filtering methods, but higher than that of only interpolation. QNR quality of CS is lower than only of interpolation, thus contradicting to visual inspection and existing experience. This is the **third argument** against it. ATWT is similar to CS, thus the same conclusions are valid.

JQM and QNR quality measure values for various methods and several parameter settings for WV-2 Munich and Melbourne data were analyzed. Conclusions derived for IKONOS Munich data seems to be valid for these two data too. For visual analysis see Figure 1. Here we have to note, that the aliasing effect is removed correctly only by HPFM. The presented investigation and results show, that the proposed JQM is well suited for quality assessment of different pan-sharpening methods.

5. CONCLUSIONS

A new Joint Quality Measure (JQM) based both on spectral and spatial quality measures is proposed, which allows a practical selection of optimal filtering parameters and comparison of different pan-sharpening methods. Experiments with a very high resolution satellite optical remote sensing data such as IKONOS and WorldView-2 data are performed. Additionally to JQM, two more joint quality measures: a simple product of spectral and spatial measures and QNR are investigated by comparing four different fusion methods: GFF, HPFM, CS and ATWT with different parameter settings.

JQM appeared to be very suitable quality measure for the selection of optimal parameters for GFF and HPFM methods, what is confirmed by visual analysis. Moreover, GFF variant (HPFM) using bilinear/cubic convolution interpolation methods outperformed all other methods on these test data showing the strength of filtering methods. Thus JQM provides a promising and practical approach to compare various pan-sharpening methods quantitatively just using a sole quality measure.

Already known joint quality measures: product and QNR seem to be not able to rank correctly different quality images, thus contradicting visual analysis. Future work will be directed towards replacement of SSIM with a recently proposed similarity measure CMSC [14].

6. REFERENCES

[1] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P.S. Huang, "A new look at HIS-like image fusion methods," *Information Fusion* 2, pp. 177-186, 2001.

[2] P.S.Jr., Chavez, S.C. Sides, and J.A. Anderson, "Comparison of Three Different Methods to Merge Multiresolution and Multispectral Data: Landsat TM and SPOT Panchromatic," *Photogrammetric Engineering & Remote Sensing* 57(3), pp. 295-303, 1991.

[3] G. Palubinskas, and P. Reinartz, "Multi-resolution, multi-sensor image fusion: general fusion framework," Proc. of Joint Urban Remote Sensing Event JURSE, 11-13 April 2011, Munich, Germany, IEEE, pp. 313-316, 2011.

[4] G. Palubinskas, "Fast, simple and good pan-sharpening method," *Journal of Applied Remote Sensing* 7(1), pp. 1-12, 2013.

[5] T. Ranchin, and L. Wald, "Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation," *Photogrammetric Engineering & Remote Sensing* 66(1), pp. 49-61, 2000.

[6] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-Driven Fusion of High Spatial and Spectral Resolution Images Based on Oversampled Multiresolution Analysis," *IEEE Transactions on Geoscience and Remote Sensing* 40(10), pp. 2300-2312, 2002.

[7] S. Klonus, and M. Ehlers, "Image fusion using the Ehlers spectral characteristics preservation algorithm," *GIScience & Remote Sensing* 44(2), pp. 93–116, 2007.

[8] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Image Processing* 13(4), 600-612, 2004.

[9] C. Padwick, M. Deskevich, F. Pacifici, and S. Smallwood, "WorldView-2 Pan-Sharpening," Proc. of ASPRS, April 26-30 2010, San Diego, California, USA, pp. 1-13, 2010.

[10] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and Panchromatic Data Fusion Assessment without Reference," *Photogrammetric Engineering & Remote Sensing* 74(2), pp. 193-200, 2008.

[11] A. Makarau, G. Palubinskas, and P. Reinartz, "Analysis and selection of pan-sharpening assessment measures," *Journal of Applied Remote Sensing* 6(1), 063548, pp. 1-20, 2012.

[12] A. Makarau, G. Palubinskas, and P. Reinartz, "Selection of numerical measures for pan-sharpening assessment," Proc. of IGARSS, Munich, Germany, IEEE, pp. 2264-2267, 2012.

[13] M. Canty, "IDL Extensions for ENVI," 2009, <http://mcanty.homepage.t-online.de/software.html> (23 May 2014).

[14] G. Palubinskas, "Mystery behind similarity measures MSE and SSIM," Proc. of ICIP, Paris, France, IEEE, 2014.

Table 1. Quality measures JQM and QNR of various pan-sharpening methods for IKONOS Munich data. In bold marked are the best values for a particular quality measure.

Method	Model	Interpolation method	f_{cutoff_HR}	QLR	QHR	JQM	f_{cutoff_HR}	D_λ	D_s	QNR
-	-	Zero padding	-	0.9548	0.6982	0.9270	-	0.0227	0.0863	0.8929
-	-	Nearest neighbor	-	0.9602	0.7727	0.9445	-	**0**	0.0585	0.9415
-	-	Bilinear	-	0.9888	0.7187	0.9481	-	0.0096	0.0767	0.9145
-	-	Cubic convolution	-	**0.9946**	0.7239	0.9520	-	0.0038	0.0755	0.9201
GFF	-	Zero padding	0.15	0.9505	0.9300	0.9708	0.45	0.0071	0.0158	0.9772
HPFM	Additive	Zero padding	0.15	0.9469	0.9237	0.9677	0.55	0.0077	0.0152	0.9773
HPFM	Additive	Nearest neighbor	0.3	0.9497	0.9011	0.9647	0.7	0.0134	0.0330	0.9541
HPFM	Additive	Bilinear	0.2	0.9755	0.8990	0.9771	0.7	0.0037	0.0135	**0.9828**
HPFM	Additive	Cubic convolution	0.25	0.9838	0.8790	0.9773	0.7	0.0082	**0.0125**	0.9794
HPFM	Multiplica	Zero padding	0.15	0.9502	0.9250	0.9697	0.5	0.0055	0.0218	0.9728
HPFM	Multiplica	Nearest neighbor	0.25	0.9459	0.9124	0.9650	0.7	0.0156	0.0347	0.9502
HPFM	Multiplica	Bilinear	0.2	0.9779	0.9015	0.9789	0.7	0.0063	0.0187	0.9751
HPFM	Multiplica	Cubic convolution	0.2	0.9808	0.8965	**0.9793**	0.7	0.0106	0.0169	0.9727
CS	Additive	Zero padding	-	0.9048	**0.9763**	0.9571	-	0.0990	0.1717	0.7464
CS	Additive	Nearest neighbor	-	0.9052	0.9742	0.9569	-	0.0862	0.1601	0.7675
CS	Additive	Bilinear	-	0.9172	0.9715	0.9630	-	0.0883	0.1641	0.7621
CS	Additive	Cubic convolution	-	0.9198	0.9736	0.9641	-	0.0829	0.1600	0.7703
CS	Multiplica	Zero padding	-	0.8916	0.9754	0.9503	-	0.0731	0.1508	0.7871
CS	Multiplica	Nearest neighbor	-	0.8960	0.9694	0.9514	-	0.0535	0.1314	0.8221
CS	Multiplica	Bilinear	-	0.8959	0.9733	0.9520	-	0.0639	0.1417	0.8035
CS	Multiplica	Cubic convolution	-	0.8954	0.9719	0.9515	-	0.0593	0.1365	0.8123
ATWT	-	-	-	0.9146	0.9554	0.9579	-	0.0728	0.1290	0.8076

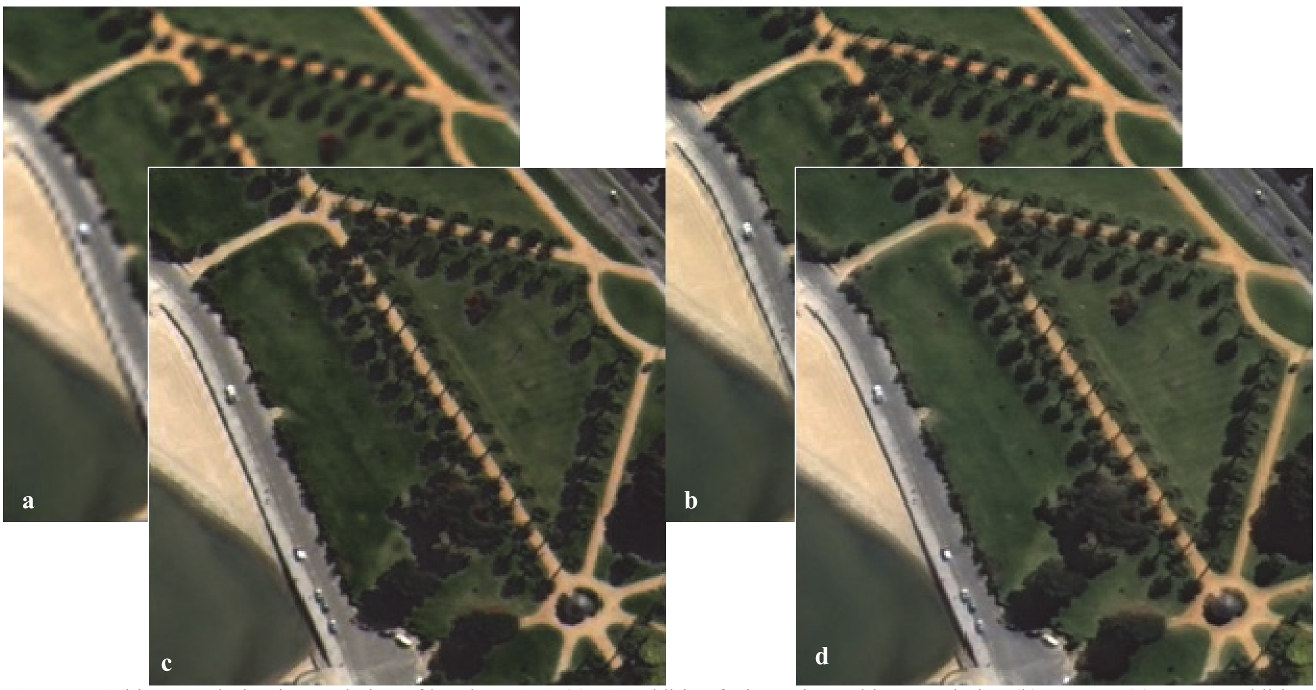

Figure 1. Cubic convolution interpolation of bands: 5, 3, 2 (a), CS additive fusion using cubic convolution (b), ATWT (c), HPFM additive fusion using cubic convolution (d) for WV-2 Melbourne data.