EVALUATING THE POTENTIAL OF IMAGE QUALITY METRICS FOR QUALITY ASSESSMENT OF HIGH RESOLUTION PAN-SHARPEN SATELLITE IMAGERY IN URBAN AREA

F. Samadzadegan, F. DadrasJavan*

Department of Geomatics Engineering, University College of Engineering, University of Tehran, Tehran, Iran – (samadz, fdadrasjavan)@ut.ac.ir

Commission I, WG I/4

KEY WORDS: Fusion, Quality Metrics, Registration Accuracy, Quantitative Analysis, Satellite Imagery

ABSTRACT:

Recently, variety of fusion methods has been proposed to enhance the spatial and spectral resolution of high resolution satellite imagery. Most of the available satellite images are restricted to improved spatial resolution or spectral resolution. Considering the direct effects of registration accuracy on the quality of fused image, it is necessary to evaluate these effects before using them in latter applications. Accordingly lots of quality evaluation processes have been proposed for quality assessment of fused images which are mostly inspired from developed image quality approaches. This paper deals with potential evaluation of some common quantitative approaches inspecting quality of fusion. Experiments conducted to evaluate the sensitivity of them to registration accuracy on Quick Bird high resolution satellite imagery in an urban area. The obtained results clearly reveal that these metrics sometimes do not behave robust and their obtained results are also inconsistence in different patch areas with different level of spectral distortion.

1. INTRODUCTION

Topographic earth observation satellites, such as IKONOS, Quick Bird and GeoEye, provide both panchromatic images at a higher spatial resolution and multi-spectral images at a lower spatial resolution but rich spectral information (Kitaw, 2007). It is due to several technological limitations for having a sensor with high spatial and spectral characteristics. So the remote sensing community has switched to merge multi-spectral and panchromatic images to exhibit complementary characteristics of spatial and spectral resolutions (Reys et al. 2004). This new product is entitled as pan-sharpened images (Ranchin and Wald, 2000). Pan sharpening has become very important in many applications of remote sensing like land use classification, detecting changes, updating maps, monitoring hazards and many other Geo-information applications (Reys et al. 2004; Ehlers et al., 2008; Kitaw, 2007).

Registration of reference images is a crucial step in image fusion (Blanc et al. 1998). Errors of co-registration quality of reference images introduce local errors in merging process and results in significant color distortions in the fused image due to the registration accuracy. So, the quality assessment of these data is crucial before using them in other next process of object extraction or recognition (Ehlers et al., 2008; Blanc et al., 1998).

Many image quality assessment algorithms have been shown to behave consistently when applied to distorted images created from the same original image, using the same type of radiometric and spectral characteristics. However, the effectiveness of these models degrades significantly when applied to a set of images originating from different reference images, and/or including a variety of different types of distortion. Considering the fact that how well an algorithm performs is defined by how well it correlates with human perception of quality, this study focuses on capability evaluation of different quantitative Image Fusion Quality Metrics (IFQMs) in comparison with qualitative quality assessment of processed images. The mentioned strategies are developed to inspect the quality of Pan-sharpening QuickBird panchromatic and multi spectral images in an urban region that enjoys variety of manmade and natural patterns.

2. IMAGE FUSION QUALITY ASSESSMENT

Image fusion quality evaluation approaches are included into two main categories as qualitative and quantitative evaluation approaches. In qualitative approach, quantifying image quality is through subjective evaluation done by human beings (Wang et al., 2004). Since this process is a time consuming process and needs expert operators, there is a wide range of research in direction of the quantitative evaluation which is based on objective performance assessment of fusion process

* Corresponding author

(Wang et al., 2004). A quantitative approach should measure the ability of fusion process to transfer all perceptually important information of input images into the output image as accurately as possible. However, quantitative performance assessment is a difficult issue due to the variety of different application requirements and the lack of a clearly defined ground-truth. A wide range of quantitative fusion assessment techniques is based on the initial concepts of image quality metrics (such as Entropy, DIV, UQI and C.C) which are already used to compare quality of two different images in image processing applications.

2.1 Qualitative analysis

The most reliable judgment of image quality assessment is subjective rating by human observer which is known as qualitative analysis (Zhang, 2006). Qualitative analysis involves visual comparison of color between original Multi Spectral and fused images, and the spatial detail between original Panchromatic and fused images (Zhang, 2008).

This method depends on the observers' experiences or bias thus some uncertainty is involved. Qualitative measure cannot be represented by rigorous mathematical models, and their techniques are mainly visual and time consuming procedures (Shi, 2005).

2.2 Quantitative analysis

Considering the draw backs of the subjective quality assessment method, much effort has been devoted to develop objective image fusion quality assessment methods (Wang et al. 2002b; Shi, 2005). Quantitative approaches involve a set of predefined quality indicators for measuring the spectral and spatial similarities between the fused image and the original Multi Spectral and/or Panchromatic images (Zhang, 2008). Amongst all developed objective quality metrics, Entropy, DIV, UQI and C.C are some of the most widely applied metrics (Riyahi et al., 2009; Wald, 2000, Thomas and Wald, 2006b). In the following a brief review on theoretical concept of these metrics is presented.

Entropy: Entropy is a measure of information content of an image and is usually applied in image processing methods as a mean for measuring the information and complexity of images (Leung et al., 2001; Sadjadi, 2005).

The Entropy of an image can be calculated by:

$$Entropy = -\sum p_i \cdot \log_2 p_i \qquad p_i = sum(image == i) / N \tag{1}$$

Where p is the estimated probability density function (normalized pixel intensity histogram) of the selected image region (Sadjadi, 2005).

For evaluating the quality of image fusion, the change in Entropy index is applied as quality metric. So we considered the change in Entropy index of each band of images before and after fusion as a metric for quality control:

$$R_E = Entropy_{Fuse\dim age} - Entropy_{initia\lim age} \tag{2}$$

It is obvious that when no change occurs in information content of images or both input images (initial and fused image) are the same, the Entropy index R_E is equal to 0.

DIV: DIV inspects fusion quality over the whole image which means difference in variances relative to the original one (Equation. 3).

$$DIV = \frac{\sigma^2_{MS} - \sigma^2_{FMS}}{\sigma^2_{MS}} \tag{3}$$

Where σ^2_{MS} is the variance of the original image and σ^2_{FMS} is the variance of the fused image. This index presents the decrease or increase of information content during fusion process and would be positive for decreasing and negative for increasing change of information.

UQI: Structural Similarity Image Metric (SSIM) introduced in (Thomas and Wald, 2006a), and more formally distilled in (Wang et al., 2004). The basic form of SSIM is very easy to understand. Suppose that x and y are local image patches taken from the same location of two images that are being compared. The local SSIM index measures the similarities of three elements of the image patches: the similarity $l(x, y)$ of the local patch luminance (brightness values), the similarity $c(x, y)$ of the local patch contrasts, and the similarity $s(x, y)$ of the local patch structures. These local similarities are expressed as Equation. 4 (Wang and Bovik 2009).

$$S(x,y) = l(x,y).C(x,y).S(x,y) = \frac{2.\bar{x}.\bar{y} + C_1}{(\bar{x}^2 + \bar{y}^2 + C_1)} \cdot \frac{2.\alpha_x.\alpha_y + C_2}{\left(\alpha_x^2 + \alpha_y^2 + C_2\right)} \cdot \frac{\alpha_{xy} + C_3}{\alpha_x.\alpha_y + C_3} \tag{4}$$

Where $\bar{x}$ and $\bar{y}$ are the local sample means of x and y, σ_x and σ_y are the local sample standard deviations of x and y, and σ_{xy} is the sample cross correlation of x and y after removing their means. The items C_1, C_2, and C_3 are small positive constants that stabilize each term. The Universal Quality Index (UQI) corresponds to the case that $C_1 = C_2 = C_3 = 0$ (Wang, et al., 2004).

$$Q = \frac{4.\alpha_{xy}.\bar{x}.\bar{y}}{(\alpha_x^2 + \alpha_y^2).(\bar{x}^2 + \bar{y}^2)} \tag{5}$$

Q index is bounded in [-1,1] and its maximum value $Q=1$ achieved when $x=y$. In this study Q index is computed locally using a sliding window moves through the images. Q index of the whole image is computed by averaging the achieved local quality indices over local regions.

$$Q = \frac{1}{N}\sum_{w=1}^{N} Q_W \tag{6}$$

Where Q_w indicates the calculated quality index within the sliding window w, and N is the total number of patches used to calculate Q index.

C.C: Correlation coefficient quantifies the closeness between two images. The correlation coefficient is computed using the following Equation:

$$C.C = \frac{\sum_1^N \sum_1^M (x-\bar{x}).(y-\bar{y})}{\sqrt{\sum_1^N \sum_1^M (x-\bar{x})^2 \sum_1^N \sum_1^M .(y-\bar{y})^2}} \tag{7}$$

The correlation coefficient value ranges from -1 to 1, where the value +1 indicates that two images are highly correlated and are very close to each other. The value -1 indicates that the images are exactly opposite to each other.

3. EXPERIMENT AND RESULTS

Robustness of mentioned quality metrics with respect to registration accuracy in comparison with visual evaluation assessed on a high-resolution QuickBird image data over an urban area that poses different paternal behavior. The original panchromatic QuickBird has 0.61m pixel while the original multispectral image has 2.4m pixel spatial resolution (for more information visit digital globe website). Applying PCI software a fused QuickBird image generated with 0.61 meter spatial resolution and three B1, B2, B3 (R,G,B) bands (Figure. 1).

For evaluating the robustness of different objective quality metrics, another fused image generated after introducing 1 pixel shift to the reference image. Since the goal of paper is comparing the capability of objective quality metrics with subjective analysis approach, 4 patches with different spectral and textural characteristics are selected in each of which different levels of color degradation were visually recognized (Figure. 1). As it is clear from Figure. 1, registration error, generates color distortion in different patches of generated Pan-sharpen image.

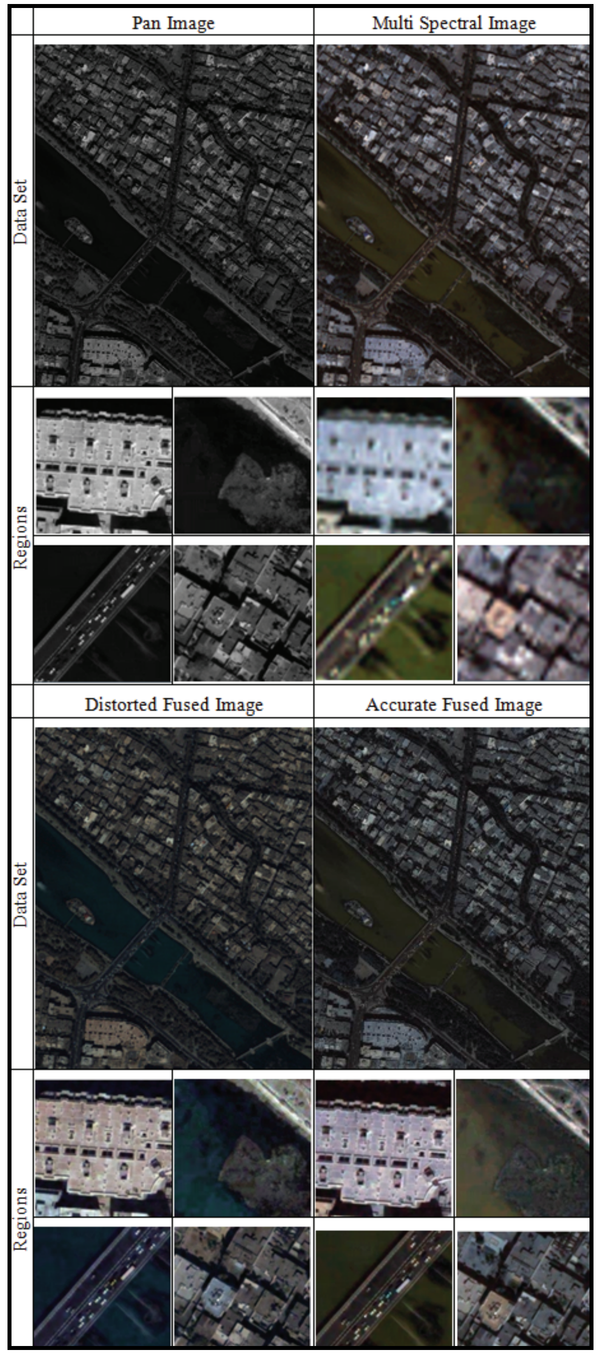

Fig.1. Quick Bird Panchromatic, Multi spectral and pan sharpen images and selected regions, based on both datasets.

Table 1. presents the overall output of objective metrics on the accurate and distorted data sets in situations of the all 4 selected patches. These metrics are Entropy, DIV, UQI and C.C.

Table. 1. Objective metrics results.

	Distance	R1	R 2	R 3	R 4
Accurate Data set	ENTROPY	0.0	-0.53	-1.14	0.23
	DIV	0.22	0.32	-0.08	0.07
	UQI	0.68	0.56	0.65	0.74
	C.C	0.96	0.85	0.86	0.93
Distorted Data set	ENTROPY	0.29	-0.22	-0.58	0.44
	DIV	-0.22	-0.51	0.94	-0.23
	UQI	0.56	0.42	0.48	0.53
	C.C	0.91	0.84	0.78	0.76

As it can be concluded from Table 1, all of IFQMs have some level of sensitivity to registration accuracy in selected patches.

3.1 Sensitivity assessment of IFQMs

In the following, discussion about capabilities and robustness of IFQMs is presented with respect to accurate and distorted image data sets.

Entropy. As it is demonstrated in Figure. 2, this metric could be considered as a good indicator for presenting the impact of the accuracy of multi spectral and panchromatic images on the fused image. However this metric suffers from a poor sensitivity to color distortion due to registration error. Besides, this metric has a weak sensitivity to local degradation of mentioned color distortion.

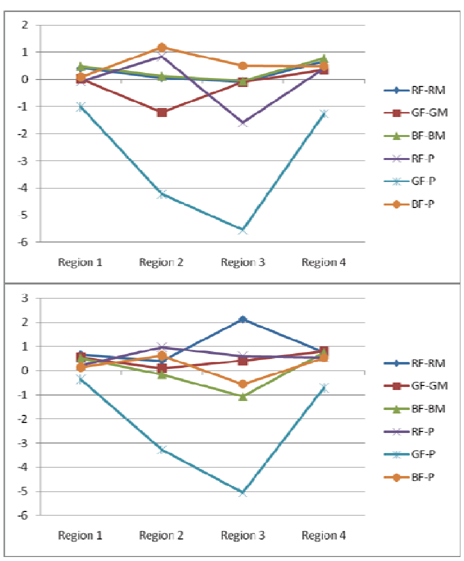

Fig.2. Entropy results for accurate data set (above) and distorted data set (below).

DIV. This metric could be considered as a good indicator for presenting the impact of reference multi spectral and panchromatic images on the fused image. Besides, it presents a good global sensitivity to color distortion due to registration error. However, it has almost a poor sensitivity to local degradation of mentioned color distortion (Figure 3).

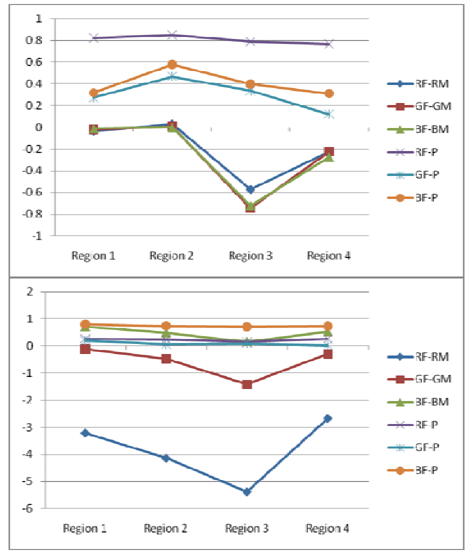

Fig.3. DIV results for accurate data set (above) and distorted data set (below).

UQI. Diverse response of UQI in terms of different bands can be considered as a characterized clue of spectral degradations (Figure 4). However, this metric suffers from averaging limitations. Besides it shows a pseudo high similarity respect to reference panchromatic band in some regions and bands and almost poor sensitivity to local color distortion (Figure 4).

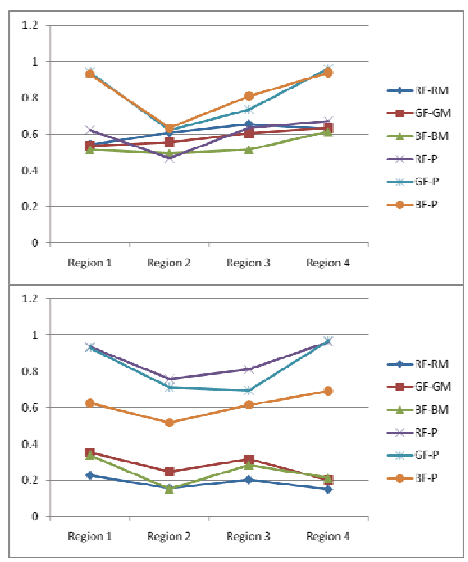

Fig.4. UQI results for accurate data set (above) and distorted data set (below).

CC. This metric has a diverse response with respect to different multi spectral and panchromatic bands which can be considered as a clue of spectral degradations. Nevertheless, it has pseudo high similarity respect to reference panchromatic band in some regions and bands (Figure. 5).

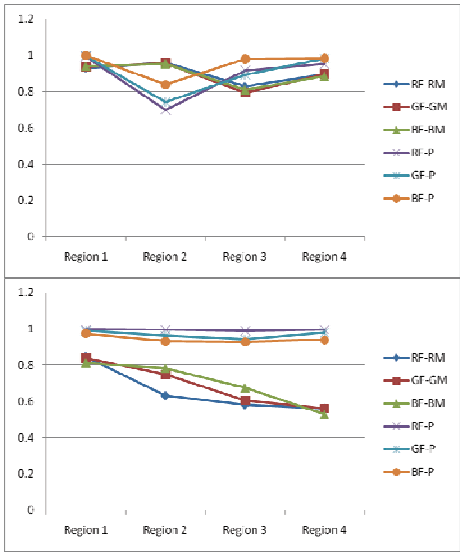

Fig.5. C.C results for accurate data set (above) and distorted data set (below).

4. CONCLUSION

Considering the importance of fusion in high resolution satellite imagery, wide range of objective image fusion quality metrics have been proposed and developed in literature. These metrics have been used in different application of remote sensing such as map production, DSM generation and urban planning. The image registration process is one of the main steps in all of image fusion techniques. This paper presented the sensitivity of image fusion quality metrics. Achieved results revealed that, although most of these metrics have acceptable capability and robustness for quantification of visual image fusion quality, some of them have a serious problem in assessments of image fusion quality under registration error. These limitations could be summarized as poor or non robust sensitivity to local degradation of colors which are clearly visually detectable and presenting pseudo high similarity respect to reference images.

An interesting direction for further work can be developing an object wise image quality metrics that could formulate the behavior of fused image based on the spectral and spatial characteristics of objects.

REFERENCES

Blanc, P., L. Wald and T. Ranchin, 1998. Importance and effect of co-registration quality in an example of « pixel to pixel » fusion process. In: Proceedings of the International Conference « Fusion of Earth data: merging point measurements, raster maps and remotely sensed images », Sophia Antipolis, France, 28-30 January.

Carvalho, Jr, O. A., Carvalho, A. P. F., Mensese, P. R., 2000, Spectral Correlation Mapper (Scm): An Improvement On The Spectral Angle Mapper (Sam). Summaries of the ninth Annual jpl airborne earth science workshop, jpl pub 00-18, jet propulsion laboratory, Pasadena, CA.

Ehlers, M. Klonus, S. Åstrand, P.J. , Quality assessment for multi-sensor multi-date image fusion. In: Proc. XXIth Int. Congr. ISPRS, Beijing, China, pp. 499–506 (2008)

Kitaw, H. G., 2007. Image pan - sharpening with Markov random field and simulated annealing. MSc Thesis, ITC, Enscede.

Leung, L.W. King, B. Vohora, V., Comparison of image data fusion techniques using entropy and INI, in proc. Acrs, vol. 1, 2001, pp. 152-157.

QuickBird spacecraft information and specifications, http://www.digitalglobe.com/index.php/85/ QuickBird

Ranchin, T., and Wald, L., 2000, Comparison of different algorithms for the improvement of the spatial resolution of images, In Proceedings of the 24th EARSeL Symposium " Fusion of Earth data: merging point measurements, raster maps and remotely sensed image", 2000, Sophia Antipolis, France.

Reyes, R.A., Gutierrez, M.J., Fernandez, S. , Thomas, C. Ranchin, T. and Wald, L. ,2004. Evaluation of the quality of Quickbird fused products. Proceedings of the 24th symposium of the European Association of Remote Sensing laboratories, Dubrovnik, Croatia, 25-27 May 2004 in New strategies for European Remote Sensing, Oluic (ed) 2005.

Riyahi, R. Kleinn C. and Fuchs, H. , Comparison of different image fusion techniques for individual tree crown identification using QuickBird images, In proceeding of ISPRS Hannover Workshop 2009,

Sadjadi, F., Comparative image fusion analysis, computer vision and pattern recognition, 2005 IEEE computer society conference on 25june 2005 page(s): 8 - 8

Shapiro, T. Barloon, P. J. and Goetz, A. F. H., "the spectral image- Processing system (sips) - interactive visualization and analysis Of imaging spectrometer data," remote sensing of environment, Vol. 44, pp. 145-163, 1993.

Shi, W., Zhu, Ch., Tian, Y. and Nichol, J., 2005. Wavelet-based image fusion and quality assessment. International Journal of Applied Earth Observation and Geoinformation. v6. 241-251.

Thomas, C. and Wald, L. , 2006a. Comparing distances for quality assessment of fused products, In Proceedings of the 26th EARSeL Symposium "New

Strategies for European Remote Sensing", 29-31 May 2006, Varsovie, Pologne.

Thomas, C. and Wald, L., 2006b. Analysis of changes in quality assessment with scales, In Proceedings of FUSION06, 10-13 July 2006, Florence, Italy.

Van Der Meer, F. , "The Effectiveness Of Spectral Similarity Measures For The Analysis Of Hyperspectral Imagery", international journal of Applied earth observation and geoinformation, vol. 93, 1-15, 2005.

Wald, L. "Quality of High Resolution Synthesized Images: Is There a Simple Criterion?" Proc. Int. Conf. Fusion Earth Data. 2000.

Wang, Z., Bovik, A. C., and Lu, L., 2002, "Why is image quality assessment so difficult?" IEEE International Conference on Acoustics, Speech, & Signal Processing, May 2002.

Wang, Z., Bovik, A.C. Sheikh, H.R. and Simoncelli, E.P. "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process., vol.13, no.4, pp.600–612, 2004.

Wang, Z. Bovik, A.C., 2009, Mean squared error: love it or leave it? IEEE Signal Processing Magazine, 2009, {26} (1): 98-117

Zhang, Y., 2008, methods for image fusion quality assessment-a review, comparison and analysis, the international archives of photogrammety, remote sensing and spatial information sciences, Vol XXXVII,(B7). Beijing

Zhang, D., 2006. Information theoretic criteria for image quality assessment based on natural science statistics, MSc Thesis, University of Waterloo.