MULTIRESOLUTION IMAGE FUSION: PHASE CONGRUENCY FOR SPATIAL CONSISTENCY ASSESSMENT

A. Makarau a, *, G. Palubinskas a, P. Reinartz b

a DLR, German Aerospace Center, 82334 Weßling, Germany - (aliaksei.makarau, gintautas.palubinskas, peter.reinartz)@dlr.de

Commission VII, WG VII/6

KEY WORDS: pan-sharpening, multispectral image fusion, spatial consistency assessment, phase congruency

ABSTRACT:

Multiresolution and multispectral image fusion (pan-sharpening) requires proper assessment of spectral consistency but also spatial consistency. Many fusion methods resulting in perfect spectral consistency may leak spatial consistency and vice versa, therefore a proper assessment of both spectral and spatial consistency is required. Up to now, only a few approaches were proposed for spatial consistency assessment using edge map comparison, calculated by gradient-like methods (Sobel or Laplace operators). Since image fusion may change intensity and contrast of the objects in the fused image, gradient methods may give disagreeing edge maps of the fused and reference (panchromatic) image. Unfortunately, this may lead to wrong conclusions on spatial consistency. In this paper we propose to use phase congruency for spatial consistency assessment. This measure is invariant to intensity and contrast change and allows to assess spatial consistency of fused image in multiscale way. Several assessment tests on IKONOS data allowed to compare known assessment measures and the measure based on phase congruency. It is shown that phase congruency measure has common trend with other widely used assessment measures and allows to obtain confident assessment of spatial consistency.

1. INTRODUCTION

Pan-sharpened data have many areas of application and therefore different requirements are posed on the fusion method. The requirements can be on spectral consistency, spatial consistency or on the both together. Spectral consistency assumes that pansharpened data have increased spatial resolution with spectral properties of the original data. Spatial consistency assumes that “A high spatial quality merged image is the one” (González-Audícana, 2005). Spectral and spatial consistency together is the ideal case of the fused data and the ideal fusion method is to provide these characteristics. A pansharpening method may provide perfect spectral consistency together with poor spatial consistency and vice versa. Therefore, to make proper decision on a fusion method (or to outline the best one), assessment of both spectral and spatial consistency is to be performed.

2. PAN-SHARPENED DATA QUALITY

2.1 Spectral consistency

Spectral consistency assessment usually performed using Wald's protocol in order to make reference multispectral data of high resolution available. There is a variety of developed and well-known similarity measures used for spectral consistency assessment. The most known and popular are: Spectral Angle Mapper, SAM (calculated as the angle between two vectors):

\[ \theta(r, f) = \cos^{-1} \left( \frac{\sum_{i=1}^{K} f_i \mathbf{r}_i \mathbf{f}_i}{\sqrt{\sum_{i=1}^{K} f_i^2} \sqrt{\sum_{i=1}^{K} (\mathbf{r}_i \mathbf{f}_i)^2}} \right), \]  

where \( K \) is the number of bands, \( r \) and \( f \) are the two vectors created by the values of spectral channels at the same pixel in the reference and fused images \( A \) and \( B \); Structural SIMilarity SSIM (Wang, 2004) or extended SSIM - Q4 (Alparone, 2003), (correlation, contrast, and luminance similarity between two images are used to calculate one similarity value):

\[ \text{SSIM}(A_i, B_i) = \left( \frac{2 \mu_A \mu_B + C_1}{\mu_A^2 + \mu_B^2 + C_2} \right) \left( \frac{2 \sigma_{AB} + C_3}{\sigma_A^2 + \sigma_B^2 + C_3} \right) \]  

where \( \mu_A \) and \( \mu_B \) are the local sample means of \( A \) and \( B \), respectively, \( \sigma_A \) and \( \sigma_B \) are the local sample standard deviations of \( A \) and \( B \), respectively, and \( \sigma_{AB} \) is the sample cross correlation of \( A \) and \( B \) after removing their means. The items \( C_1 \), \( C_2 \), and \( C_3 \) are small positive constants that stabilize each term; ERGAS (Wald, 1997) (similarity measure for multispectral images, based on MSE estimator):

\[ \text{ERGAS}(A_i, B_i) = 100 \frac{h}{l} \frac{1}{K} \sum_{i=1}^{K} \frac{\text{RMSE}(A_i, B_i)^2}{\mu_A^2}, \]  

* Corresponding author. Tel. +49-8153-283672
where $A_i$ and $B_i$ are the compared bands of a multispectral image, RMSE is root mean squared error, $\mu_A$ is the mean value of $A_i$, $K$ is the number of bands, $h$ is high/low resolution images ratio; Zero mean normalized cross-correlation, ZNCC:

$$
ZNCC(A_i, B_i) = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} (A_i(m,n) - \mu_{A_i})(B_i(m,n) - \mu_{B_i})}{\sqrt{\sum_{m=1}^{M} \sum_{n=1}^{N} (A_i(m,n) - \mu_{A_i})^2} \sqrt{\sum_{m=1}^{M} \sum_{n=1}^{N} (B_i(m,n) - \mu_{B_i})^2}}
$$

where $A_i$, $B_i$ are the compared images; $\mu_{A_i}, \mu_{B_i}$ are the averages of the images $A_i, B_i$, respectively; $M, N$ is the size of the compared images.

### 2.2 Spatial consistency

Spatial consistency is another aspect of fused imagery assessment. Up to now not many papers deal with spatial consistency assessment. Almost all the works use single scale edge detector (Gradient, Laplacian, Sobel edge detector) and an evaluation metric to calculate the distance between the edge maps (usually correlation coefficient) (Shi, 2003; Zhou, 1998; Pradhan, 2006). Here the comparison is made between the fused bands and the corresponding panchromatic image. Another approach calculates the percentage of true and false edges introduced into the fused band using Sobel edge detector (Pradhan, 2006). Several works on fusion report use of SSIM and ERGAS measures for spatial consistency assessment (Lillo-Saavedra, 2005) (panchromatic image is used as the reference instead of a spectral band).

In this paper we propose to use an additional measure for spatial consistency assessment. This measure uses phase congruency (PC) (Kovesi, 1999) for feature extraction on an image. Invariance to intensity and contrast change as well as multiscale nature of this measure allows to obtain more confident assessment comparing to single-scale edge detectors.

### 3. PHASE CONGRUENCY FOR SPATIAL CONSISTENCY ASSESSMENT

#### 3.1 Phase congruency

Phase congruency was proposed as intensity and contrast invariant dimensionless measure of feature significance, and used for signal matching and feature extraction (Kovesi, 1999). Phase congruency at point $x$ may be defined in the following way:

$$
PC(x) = \frac{\sum_{\alpha} W_{\alpha}(x) | F_{A_{30}}(x) \Delta \Phi_{A_{30}}(x) - T_o |}{\sum_{\alpha} \sum_{\alpha} F_{A_{30}}(x) + \varepsilon},
$$

where $F_{A_{30}}$ is the amplitude of the component in Fourier series expansion, $\Delta \Phi_{A_{30}}$ is the phase deviation function, $W_{\alpha}$ is the PC weighting function, $o$ is the index over orientation, $s$ is the index over scale, $T_o$ is the noise compensation term, $\varepsilon$ is the term added to prevent division by zero, $\lceil \rceil$ means that the enclosed quantity is permitted to be non-negative (Kovesi, 1999).

A bank of 2D Log Gabor wavelets is used for feature extraction. Different scale and orientation of the wavelets in the bank allow extracting more information about the structure (detail) of the image under assessment.

Multiscale image analysis instead of single-scale gradient operators allows to extract more information on image structure, features and edges. The result of PC extraction is phase congruency feature map. This map represents the structure of the image and allows to perform feature based image comparison.

#### 3.2 Comparison metric

Zero mean normalized cross correlation was selected as a comparison metric of PC feature maps. Liu et. al. report on successful application of the metric for this task (Liu, 2008). ZNCC produces a real value in the range [-1,1] where 1 indicates full similarity of compared maps and -1 indicates absolute dissimilarity.

Pan-sharpened spectral band and corresponding panchromatic image are used for extraction of PC feature maps, and the maps are compared using ZNCC. The panchromatic image is used as the reference image for spatial consistency assessment (Figure 1).

#### 3.3 Assessment protocol

The benefit of PC application for assessment may be illustrated by comparison with other assessment methods on pan-sharpened dataset, which consists of fused images with known quality. PC is expected to show similar trend with other assessment measures and provide similar assessment results. Well-known fusion methods should be used in order to produce the dataset with expected quality.

Several well-known pan-sharpening methods were selected to produce fused images with expected quality (spatial and spectral consistency); Intensity-Hue-Saturation (IHS) image fusion (Welch, 1987), image fusion using Principal Component Analysis (PCA) (Welch, 1987), wavelet image fusion (Aiazzi, 2002), and General Image Fusion method (GIF) (Wang, 2005). Generally, well-known methods IHS and PCA produce fusion results with proper spatial consistency; wavelet fusion produces proper spectral consistency; GIF method produces a compromise of acceptable spectral and spatial consistency. Fusion methods can be sorted according to the quality of the produced result: in the sense of spectral or in the sense of spatial consistency. These methods were chosen as reference methods to produce expected results for pan-sharpened dataset used for assessment and comparison.

During the first assessment setup, a set of multispectral images was pan-sharpened by the following methods: IHS, PCA, Ah trous wavelet image fusion (ATWT, cubic B-spline), and by two modifications of General Image Fusion method (GIF-1 and GIF-2). GIF-1 extracts high-resolution image detail (high frequency component) from panchromatic image and adds to interpolated spectral image. The amount of transferred image detail data is established using regression (Starovoitov, 2007). GIF-2 employs image detail addition to interpolated spectral image (Ehlers, 2004). IHS and PCA image fusion methods were run using ENVI software, while all the other fusion methods were implemented using IDL system.
One example of quantitative assessment of IKONOS urban subscene (Athens, panchromatic image size is 4000x4000) is presented in Table 1: spectral consistency (SSIM, ERGAS, SAM) and spatial consistency assessment (SSIM PAN, ERGAS PAN, CORR PAN, HPCC, PC ZNCC). The SSIM PAN, ERGAS PAN, CORR PAN are the measures notation for spatial consistency assessment (fused image is compared with corresponding panchromatic image). Mean values of the measures are calculated over the assessed spectral channels.

The results of quantitative assessment during the second assessment setup are presented in Table 2. The dependencies of the measures on the quality of the fused images are presented in Figures 2 and 3. The characteristics of the resulting images are dependent on the GIF-2 hf parameter. Assessment of the pansharpened images with different quality (GIF-2 method, parameter variation) results in different scores and allows to illustrate trends of the measures.

4. RESULTS AND DISCUSSION

One of important questions during this investigation was: does the assessment using PC has the same trend with the other measures? The results produced by assessment measures were analysed for similarity in trend.

The PC ZNCC and SSIM PAN, ERGAS PAN, CORR PAN, HPCC illustrate higher spatial consistency produced by the IHS, PCA, GIF-1 and GIF-2 methods. This agrees with the well-known fact that the IHS, PCA and GIF methods produce the best spatially-consistent results with some loss of spectral consistency. Here the PC ZNCC illustrates similar results comparing with other measures on spatial consistency (Table 1). For the ATWT fusion, the PC ZNCC and SSIM PAN, ERGAS PAN, CORR, HPCC illustrate loss of spatial consistency and the highest spectral consistency (SSIM and ERGAS). PC-based metric resulted in the lowest value on spatial consistency, which correlates with the knowledge about the fusion result. GIF-1 and GIF-2 methods provided a compromise between the spectral (SSIM, SAM) and spatial consistency (PC ZNCC and SSIM PAN, ERGAS PAN, HPCC, together).

Highest score of SAM for GIF-2 method (Table 1) was caused by characteristics of the General Image Fusion (GIF) method, which provides a good compromise between the spatial and spectral consistency. For this particular case, the GIF method resulted in good spectral consistency (according to SAM measure) with acceptable spatial consistency.

Table 1 illustrates better values of ERGAS PAN for ATWT (3.78) than for IHS (11.17). The opposite trend is shown by the PC ZNCC and CORR PAN, HPCC. Such results may originate from instability of the MSE estimator (Wang, 2009) in ERGAS measure. Also SSIM PAN illustrated low spatial consistency of the IHS fusion (SSIM PAN (mean) equals to 0.6314). This
disagreement may be caused by the nature of SSIM measure, which uses comparison of luminance and contrast of the images. For this example, PC ZNCC assessment is not skewed and coincides with results of HPCC and Correlation.

The second assessment setup is expected to illustrate increase of spectral consistency with simultaneous decrease of spatial consistency on the fusion results (GIF-2 method, change of parameter set). Dependency graphs of the assessment measures are presented in Figure 2 (spectral consistency: SSIM mean, ERGAS mean, SAM) and in Figure 3 (spatial consistency: SSIM PAN (mean), ERGAS PAN, CORR PAN (mean), HPCC (mean), PC ZNCC (mean)). Since the ideal values for SSIM, ERGAS, and SAM are (respectively) 1, 0, and 0, the spectral consistency measures are increasing (Figure 2). This corresponds to our assumption and expectation. For the spatial consistency assessment, the ideal values for SSIM PAN, ERGAS PAN, CORR PAN, HPCC, PC ZNCC are (respectively) 1, 0, 1, and 1, the spatial consistency measures are decreasing (Figure 3). This also corresponds to our assumption and expectation for spatial consistency. Figure 3 clearly illustrates similar trend of the PC ZNCC with all the other spatial consistency measures.

Table 2 illustrates common trend on spatial consistency between the results obtained by known measures and the PC-based metric. Spatial consistency assessment using PC illustrates expected decrease of spatial consistency. Also, the PC ZNCC measure is more sensitive to change of spatial consistency, so it is easier to assess and compare the quality of the image.

Visual assessment shows that the best spatial consistency have the IHS, PCA, GIF-1, and GIF-2 methods while ATWT resulted in slightly blurred edges (Figure 4), and coincides with the results of numerical assessment using PC. Figure 5 presents corresponding fragments of panchromatic and fused image (IHS fusion), edge maps (Sobel operator), and maximum moment of PC covariance (indicator of edge strength). It should be noted that PC feature map should not be confused with edge map. Figure 5 illustrates difference of image intensity and contrast (subfigures a, b)). Different edge maps are produced by edge detection operators (subfigures c, d). It is also demonstrated that the PC is more stable to intensity and contrast change (subfigures e, f).

5. CONCLUSIONS

Not many papers report on spatial consistency assessment of pan-sharpened data. Therefore, a need for robust and sufficient measures still exists. Application of phase congruency for spatial consistency assessment is proposed. Multiscale nature of phase congruency as well as invariance to intensity and contrast change allows more thorough analysis of fused data, comparing to single-scale edge detection methods. Identical trend with different assessment measures and with visual assessment showed that phase congruency is relevant for spatial consistency assessment, and the decision on the consistency can be made with higher confidence. Also it was found that ERGAS and SSIM provided less stability for spatial consistency assessment than correlation and edge-based measures. It should be noted that sometimes use of single assessment measure is not sufficient and may give skewed results (not all the characteristics of the fused data are revealed). Therefore use of several assessment measures increases confidence over calculated results.

REFERENCES


### Table 1. Spectral and spatial consistency assessment of the pan-sharpened image dataset (first assessment setup)

<table>
<thead>
<tr>
<th>Fusion Method</th>
<th>SSIM, ideal=1 (mean)</th>
<th>SSIM PAN, ideal=1 (mean)</th>
<th>ERGAS, ideal=0 (mean)</th>
<th>SAM, ideal=0 (mean)</th>
<th>SSIM PAN, ideal=1</th>
<th>ERGAS PAN, ideal=0 (mean)</th>
<th>CORR PAN, ideal=1 (mean)</th>
<th>HPCC, ideal=1 (mean)</th>
<th>PC ZNCC, ideal=1 (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ATWT</td>
<td>0.9527</td>
<td>0.8883</td>
<td>1.2804</td>
<td>1.0164</td>
<td>0.8639</td>
<td>0.7474</td>
<td>0.8181</td>
<td>0.8212</td>
<td>0.7488</td>
</tr>
<tr>
<td>2 HS</td>
<td>0.1737</td>
<td>0.2042</td>
<td>0.2767</td>
<td>0.2102</td>
<td>0.6184</td>
<td>0.5890</td>
<td>0.6870</td>
<td>0.6314</td>
<td>11.1713</td>
</tr>
<tr>
<td>3 PCA</td>
<td>0.8036</td>
<td>0.6736</td>
<td>0.6311</td>
<td>0.7103</td>
<td>0.9307</td>
<td>0.9043</td>
<td>0.9754</td>
<td>0.9628</td>
<td>0.9516</td>
</tr>
<tr>
<td>4 GIF-1</td>
<td>0.7462</td>
<td>0.6079</td>
<td>0.5693</td>
<td>0.6386</td>
<td>0.9457</td>
<td>0.9054</td>
<td>0.9650</td>
<td>0.9447</td>
<td>0.9447</td>
</tr>
<tr>
<td>5 GIF-2 (90%)</td>
<td>0.7057</td>
<td>0.6666</td>
<td>0.7203</td>
<td>0.7288</td>
<td>0.7076</td>
<td>0.6405</td>
<td>0.7990</td>
<td>0.9157</td>
<td>0.9216</td>
</tr>
</tbody>
</table>

### Table 2. Spectral and spatial consistency assessment of GIF-2 pan-sharpened image dataset (second assessment setup)

<table>
<thead>
<tr>
<th>Fusion Method</th>
<th>SSIM, ideal=1 (mean)</th>
<th>SSIM PAN, ideal=1 (mean)</th>
<th>ERGAS, ideal=0 (mean)</th>
<th>SAM, ideal=0 (mean)</th>
<th>SSIM PAN, ideal=1</th>
<th>ERGAS PAN, ideal=0 (mean)</th>
<th>CORR PAN, ideal=1 (mean)</th>
<th>HPCC, ideal=1 (mean)</th>
<th>PC ZNCC, ideal=1 (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 GIF-2 (90%)</td>
<td>0.7057</td>
<td>0.6666</td>
<td>0.7203</td>
<td>0.7288</td>
<td>0.7076</td>
<td>0.6405</td>
<td>0.7990</td>
<td>0.9157</td>
<td>0.9216</td>
</tr>
<tr>
<td>2 GIF-2 (75%)</td>
<td>0.7333</td>
<td>0.7011</td>
<td>0.7594</td>
<td>0.7529</td>
<td>0.7366</td>
<td>0.7366</td>
<td>0.7002</td>
<td>0.9191</td>
<td>0.9837</td>
</tr>
<tr>
<td>3 GIF-2 (50%)</td>
<td>0.8205</td>
<td>0.8053</td>
<td>0.8462</td>
<td>0.8391</td>
<td>0.8277</td>
<td>0.8277</td>
<td>0.7002</td>
<td>0.9191</td>
<td>0.9837</td>
</tr>
</tbody>
</table>

Figure 2. Dependency of spectral consistency measures on added high frequency in GIF-2 method (Table 2)

Figure 3. Dependency of spatial consistency measures on added high frequency in GIF-2 method (Table 2)
Figure 4. Example of a region taken from a fused urban scene (IKONOS, Athens, region size is 316x316). a) panchromatic image, b) interpolated multispectral (bilinear), c) ATWT, d) IHS, e) PCA, f) GIF-1, g) GIF-2. Fused images are composed of visible range spectral channels. ATWT fusion illustrates some loss of spatial consistency.

Figure 5. Region of the scene (München), edge and PC feature maps: a) Panchromatic image, b) IHS fusion (Band 3, red colour spectral range), c), d) Sobel edge map on panchromatic image and on IHS fusion, e), f) maximum moment of phase congruency covariance on panchromatic image and on IHS fusion.